

Supporting Online Material for

“Loss of function at *RAE2*, a previously unidentified EPFL, is required for awnlessness in cultivated Asian rice”

Kanako Bessho-Uehara^{1*}, Diane R. Wang^{2*}, Tomoyuki Furuta¹, Anzu Minami¹, Keisuke Nagai¹, Rico Gamuyao¹, Kenji Asano¹, Rosalyn B. Angeles-Shim¹, Yoshihiro Shimizu¹, Madoka Ayano¹, Norio Komeda¹, Kazuyuki Doi³, Kotaro Miura⁴, Yosuke Toda⁵, Toshinori Kinoshita⁵, Satohiro Okuda⁵, Tetsuya Higashiyama⁵, Mika Nomoto⁶, Yasuomi Tada⁶, Hidefumi Shinohara⁶, Yoshikatsu Matsubayashi⁶, Anthony Greenberg², Jian Zhong Wu⁷, Hideshi Yasui⁸, Atsushi Yoshimura⁸, Hitoshi Mori³, Susan R. McCouch^{2†} & Motoyuki Ashikari^{1†}

This PDF file includes

Materials and Methods

References

Supplemental Figures

Supplemental Tables

SUPPORTING ONLINE MATERIAL

Materials and Methods

Plant materials and growth conditions. For the positional cloning of RAE2, we used 8,000 F₂ plants derived from the cross between *O. sativa* ssp. *japonica* cv. Koshihikari and GLSL25, a chromosome segment substitution line (CSSL) carrying approximately 11.5 Mb of *O. glaberrima* Acc. IRGC104038 chromosome 8 fragment in Koshihikari background. The plant materials were grown in the research field of Nagoya University, Japan. For RAE2 complementation test and overexpression test, the awnless *O. sativa* ssp. *japonica* cv. Taichung65 (T65) and Nipponbare were used for transformation. For the RNAi test, the awned CSSL line GIL116 carrying a chromosome 8 fragment of IRGC104038 in T65 background (Fig. S2) was used to suppress RAE2 expression. Since Koshihikari have low regeneration ability in the mature seed culture system (41), Nipponbare and T65 were used instead. The transgenic plants were grown in isolated greenhouses under long day condition until the ten-leaf stage, and transferred to short day condition until flowering. Information on accessions used for the diversity and selective sweep analyses are listed in Table S2 and S4.

Primers. The primers used in this study are listed in Table S7.

Plasmid construction. The BAC clone OglA0006B21 harboring the entire 80 kb candidate region was screened from the CG14 BAC library and sequenced by shotgun sequencing. The BAC sequence was assembled using the GENETYX software package (GENETYX Co., Tokyo, Japan). The annotated genes were compared with gene annotations of Nipponbare (*O. sativa*) in RAP-DB (<http://rapdb.dna.affrc.go.jp/viewer/gbrowse/build4>). The BAC clone was partially digested with *Sau3AI*, yielding 10-30 kb fragments which were then sub-cloned into the binary vector TAC7. Five sub-clones that cover the entire 80 kb candidate region were selected and used for the complementation tests. Sub-clones #33 and #89 harbor the entire RAE2 gene. For complementation test of RAE2, the full cDNA sequence of RAE2 including 3 kb upstream of the start codon and 1 kb downstream of the stop codon of the gene was cloned into pENTR/D-TOPO (Invitrogen) and transferred into pGWB501 (42) through Gateway cloning technology (Invitrogen) to develop the transformation construct *pRAE2::RAE2*. For 4 types of RAE2 overexpression, we used modified pCAMBIA1380 including OsACTIN1 promoter. Four RAE2 DNA fragments were PCR-amplified from cDNA and cloned into modified pCAMBIA1380 treated with BamHI and HindIII. For making RAE2 overexpression construct fused with 3xFLAG, firstly KU71 and KU72 (Table S7) were used to make primer dimer then cloned into modified pCAMBIA1380 treated with NcoI. RAE2 DNA fragments were PCR-amplified from cDNA by KU73 and KU75 (Table S7) and cloned into modified pCAMBIA1380 fused with 3xFLAG treated with XbaI and HindIII. For RNAi silencing, 200 bp of 3'UTR in RAE2 was cloned into pENTR/D-TOPO and transferred into pANDA (43, 44) through Gateway cloning technology. The constructs were used for the standard protocol of Agrobacterium (strain EHA105)-mediated rice

transformation. Transgenic lines were selected on Murashige & Skoog medium plates containing 50 mg hygromycin (Sigma). To make the recombinant RAE2 pro-peptide construct, the RAE2 overexpression construct fused with 3xFLAG used as template and encoding the pro-RAE2 (from 91-AGG...CCC-375) without stop codon was PCR-amplified by PRO-f and PRO-r (Table S7). Pro-RAE2 amplicon cloned into pET32a(+) vector (Novagen) treated with EcoRV and XhoI. This construct used as template for making a series of muRAE2 constructs which are alanine-substituted around cleavage region made by site directed mutagenesis according to the protocol.

Phenotypic evaluation. Panicles of the parental plants (Koshihikari and each CSSL) and transgenic plants (BAC sub-clones and RAE2 genomic fragment complementation lines, RNAi lines and overexpression lines) were harvested after seed maturation. Panicles were sampled from 10 plants to measure awn length and frequency of awned seeds. Frequency of awned seeds per panicle was calculated as the number of awned seeds (5 mm<) per panicle divided by the total number of seeds per panicle. Measurements for each complementation line and RNAi line are shown in Table S1.

RNA isolation and quantitative RT-PCR. Total RNA was extracted by RNeasy Plant Mini Kit (QIAGEN), whereas first-strand cDNA synthesis was performed using the Omniscript RT Kit (QIAGEN). StepOne™ Real-Time PCR system (Applied Biosystems) was used to analyze the relative expression levels of the target genes (e.g. RAE2, rae2). Relative expression levels of the target genes were normalized to the levels of endogenous ubiquitin transcripts (OsUBI). Each set of experiments was repeated three times, and the Comparative CT method ($\Delta\Delta$ CT Method) was used to calculate the relative expression levels of the target genes. For qRT-PCR analysis of RAE2 and rae2, various plant parts (leaf blade, leaf sheath, stem, root and panicle) of Koshihikari and GLSL25 were used. Leaf blade, leaf sheath and roots were obtained from plants that were < 15 cm in height. Stems and panicles (<1 cm) were obtained from the plants 24 days after transplanting under short day condition. Error bar indicates standard deviation of the mean.

Semi quantitative RT-PCR. RT-PCR was performed in a 50 μ L solution containing a 2.5 μ L aliquot of cDNA as the DNA template, 0.2 μ M gene-specific primers (see Table S7), 10 mM deoxynucleotide triphosphates, 1 unit of ExTaq DNA polymerase (Takara), and reaction buffer. Amplifications of OsACTIN1 cDNAs was used as internal control. The reaction included an initial 5-min denaturation at 94°C, followed by 25 cycles of PCR (94°C for 30 s, 56°C for 30 s, and 72°C for 30 s), and a final 5-min extension at 72°C. The number of cycles used for amplification with each primer pair was adjusted to be in the linear range. All RT-PCR data are representative of at least three independent experiments.

Scanning electron microscopy. Developmental stages are classified into Sp7, Sp8, and post Sp8 according to Oryzabase classification (<http://www.shigen.nig.ac.jp/rice/oryzabase/devstageineachorgan/list>). The young panicles of Koshihikari and GLSL25 were fixed in starch-based glue for microscopic observation. The samples were viewed using the SEM (S-3000N, Hitachi, Tokyo, Japan) scanning electron microscope which

was set at -5°C inside temperature and at 3.2 kV.

***in situ* hybridization.** Plant materials were fixed in 4% (wt/vol) paraformaldehyde and 0.25% (vol/vol) glutaraldehyde in 0.1 M sodium phosphate buffer (pH 7.2) overnight at 4°C, dehydrated through a graded ethanol series followed by a t-butanol series, and finally embedded in Paraplast Plus (Sherwood Medical). Microtome sections (8-10 µm) were mounted on adhesive glass slides (Matsunami Glass Ind., Ltd). Digoxigenin-labeled RNA probes were transcribed with T7 RNA polymerase. The probes were amplified using the respective primer set for RAE2 and rae2 and cloned into the pBluescript II SK+ and pBluescript II KS+ vectors. Hybridization and immunological detection of the hybridized probes were performed according to a described method (45) with some modifications.

Protein extraction and immunoblot analysis. Crude protein extracts from several organs (e.g. callus, stem, leaf and spikelet) of pAct::RAE2-3xFLAG transgenic plants in the background of Nipponbare were prepared by grinding with liquid nitrogen. Total protein was extracted with 3.0 mL protein extraction buffer (20 mM Tris-HCl (pH 7.5), 1 mM EDTA, 150 mM 2% Protease inhibitor cocktail (Complete, Roche), 0.1% TritonX). After centrifugation, supernatant mixed with an equal volume of 2× sample buffer (135 mM Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 0.2 w/v % bromphenol blue, 200 mM DTT) and boiling for 5 min. Protein samples were separated by 15% SDS-PAGE and transferred to PVDF membrane (0.2 µm pore size, Millipore) by semi-dry blotting. The blots were treated with 5% skim milk in TBST (0.1w/v% Tween20, 2 mM Tris Base, 13.7 mM NaCl, pH 7.4) for 1 h and subsequently incubated with anti-FLAG antibody (1:3,000) (v/v) (A8592, Sigma) for 2 h. Blots were washed three times with TBS-T for 10 min each. Goat anti-mouse IgG horseradish peroxidase–conjugated secondary antibody was incubated for 1 h, and blots were washed following the same procedure described above. All reactions were conducted at room temperature. Detection of peroxidase activity was performed according to the instruction manual from Pierce (Thermo Fisher, Massachusetts, USA).

Purification of recombinant RAE2. The recombinant RAE2 pro-peptide construct is fused with 3xFLAG in C-terminus without stop codon for detection by FLAG antibody. The construct detail is described above. This fusion RAE2 peptide and a series of amino acid substitution-mutated peptides (muRAE2) were expressed in *E. coli* strain Rosetta (DE3) pLysS (NOVAGEN). The expressed recombinant proteins were purified by TALON beads (Clontech) according to the manufacturer's instructions. The beads were washed 5 times with a wash buffer (50 mM Tris-HCl, 100 mM NaCl, 0.1 % TritonX, 1 mM imidazole) and the recombinant proteins were collected using the elution buffer (50 mM Tris-HCl, 100 mM NaCl, 0.1 % TritonX, 10 mM imidazole). The production of recombinant peptides was confirmed by 15% SDS-PAGE.

***in vitro* processing assay.** To prepare the plant extracts, 1.0 g of each rice tissue (callus, leaf, stem, spikelet (<1 cm)) was collected and ground in liquid nitrogen following the procedure described above. The ground extract was centrifuged (15,000 rpm for 30 min at 4°C) and the resulting supernatant was used for the *in*

in vitro processing assay. For the assay, 0.5 µg of pro-RAE2-3xFLAG protein or other mutant proteins were mixed with 10 µg of each plant extract and incubated for 2 h at room temperature with or without 0.1% protease inhibitor cocktail (Complete, Roche). The concentration of recombinant peptides was determined using the Bio-Rad Protein Assay (Bio-Rad Laboratories). After incubation the peptides were separated by 15% (Fig. 4C and Fig. S10D) or 20% (Fig. 4A, 4E and S10E) SDS-PAGE. Immunoblotting was performed following the procedure described above. We used anti-FLAG antibody (1:3,000) (v/v) for Fig. 4A, 4C, 4E and S10E and anti-RAE2 antibody (1:1,000) (v/v) for Fig. S10D as primary antibody.

Specific antibody. A peptide antigen with 7-amino-acid (NH- 119 RDRLFDP 125 -COOH) in C-terminal region of RAE2 was synthesized, purified, and conjugated with keyhole limpet hemocyanin. The conjugate was injected into a rabbit to induce the production of anti-RAE2 polyclonal antibodies. These antibodies were purified from the rabbit serum using a HiTrap NHS-activated HP column (GE Healthcare) conjugated with the 7-amino-acid antigen peptide in accordance with the manufacturer's protocol.

Phylogenetic tree. RAE2 sequence was identified through reciprocal best-BLAST match searches of the Phytozome and National Center for Biotechnology Information (NCBI) databases. Accession numbers or locus IDs of EPF/EPFLs were derived from the NCBI database. Amino acid sequences for the C-terminal mature peptide region were aligned using the ClustalW program. The number of amino acid substitutions between each pair of EPF/EPFL proteins was estimated using the Jones-Taylor-Thornton (JTT) model with complete-deletion option. The phylogenetic tree was reconstructed by the neighbor-joining method. Bootstrap values were estimated (with 1000 replicates) to assess the relative support for each branch, and bootstrap values were labeled with cutoff at 50. To construct the phylogenetic tree, the neighbor-joining method in MEGA version 6.1 was used.

RAE2 3-D conformation modeling. Three-dimensional structures of RAE2 peptide were predicted by homology modeling system of the Mäestro (Schrödinger, NY, USA) software. The structure of Stomagen (protein database ID: 2LIY) determined by NMR was used as the template. Pairwise alignment was improved manually by minor editing based on the secondary structure predictions and disulfide bonds were allowed to form during the modeling. The hypothetical structure of rae2 (OsePFL1) could not be modelled because it lost two cysteine residues (C5 and C6) and did not fit the Stomagen template. Yellow: cysteine residues, red: disulfide bonds and atoms, blue: antiparallel beta sheets.

Diversity analysis. PCR products amplifying the RAE2 gene were assayed for polymorphisms using BigDye Terminator v3.1 Cycle Sequencing Kit and analyzed with CodonCode Aligner 6.0.2. All sequences were aligned to the rice reference genome (cv. Nipponbare) and predicted cDNAs were extracted and translated. SNPs and indels detected were used to construct RAE2 gene haplotypes (n=123 with full sequence). Polymorphisms with a minor allele count (MAC) <1 were filtered out for gene haplotype construction

unless they represented a frameshift mutation. The geographical map displaying origin of diverse rice accessions was created using R package 'maps.'

Selective Sweep Analysis. 67 *O. sativa* and 65 *O. rufipogon/O. nivara* accessions were analyzed for evidence of selective sweep listed in Table S4. SNP information on the *O. sativa* set were extracted from re-sequencing data (unpublished, McCouch) and imputed for RAE2 protein variant (4C, 6C, 7C) using tag SNPs from a rice SNP array identified on an overlap set of *O. sativa* that were Sanger Sequenced in the diversity analysis (46). SNP data on the *O. rufipogon/O. nivara* set were derived from Genotyping-By-Sequencing (GBS) information on chromosome 8 (unpublished, McCouch). An overlap SNP set between the re-sequencing data (1,137,573 markers on chromosome 8) and the GBS data (34,267 markers on chromosome 8) were used for estimation of nucleotide diversity and distance. π (nucleotide diversity) and d (distance between groups) statistics were calculated using sliding windows of 100 SNPs, with step size 2 variants, across chromosome 8 (47). We enumerated the sequence differences between a given pair of DNA segments and calculated sequence differentiation using the Jukes-Cantor model (48). Genetic distances between population pairs and nucleotide diversity within populations were estimated based on Nei (1973) (49). To enable comparisons between different analyses, we estimated per-kb values of π and d by dividing the total value for a window by the reference map distance (in kb) between the first and last SNP. Since only sub-sets of sites on the chromosome was covered by sequence, this procedure results in a drastic underestimate of π and d . However, the degree of underestimation is the same across groups so values are comparable within our data set.

References

41. Nishimura, A. et al., (2005) Isolation of a rice regeneration quantitative trait loci gene and its application to transformation systems. *Proc. Nat. Acad. Sci. U S A* 102(33):11940–11944.
42. Nakagawa T, et al. (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioengi.* 104(1): 34-41.
43. Miki D, Shimamoto K (2004) Simple RNAi vectors for stable and transient suppression of gene function in rice. *Plant Cell Physiol.* 45(4): 490-495.
44. Miki D, Itoh R, Shimamoto K (2005) RNA silencing of single and multiple members in a gene family of rice. *Plant Physiol.* 138(4): 1903-1913.
45. Kouchi H, Hata S (1993) Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol. Gen. Genet.* 238(1-2): 106-119.

46. McCouch S. et al. (2016) Open access resources for genome-wide association mapping in rice. *Nat. Comm.* 7:10532.
47. Weir BS, Cockerham, CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution* 38(6): 1358-1370.
48. Li WH (1997) *Mol. Evolution*, Sinauer Associates, Sunderland, Massachusetts
49. Nei, M (1973) Analysis of gene diversity in subdivided populations. *Proc. Nat. Acad. Sci. U S A* 70(12): 3321-3323.

Fig. S1. Positional cloning and complementation test of RAE2.

(A) Gross morphology of the panicles of GLSL25 (awned) and Koshihikari (awnless). F₁ derived from the cross between Koshihikari and GLS25 showed an awn phenotype, whereas the F₂ population segregated into awned and awnless at a 3:1 ratio, indicating the dominance of the RAE2 gene from *O. glaberrima*. (B) RAE2 was further delimited to an 80 kb genomic region between the markers 8KG23941 and 8KG24021. (C) Blue and red arrows represent the 12 genes within the candidate region. The orange line represents BAC clone from CG14 (*O. glaberrima*); OglA0006B21 overlapped entirely with the 80kb candidate region on chromosome 8. (D) Five sub-clones (#9, #59, #33, #89, #2-82) constructed from OglA0006B21 by partial digestion with *Sau3AI*. (E) Seed phenotypes of transgenic lines derived from each sub-clone construct in cv. Taichung65 (T65, *O. sativa* ssp. *japonica*) genetic background. Transgenic lines carrying the sub-clones #33 containing 29 kb and #89 containing 13kb fragment exhibited the awned phenotype. Line #33 had shorter awns than #89, despite having a larger region from *O. glaberrima*. This suggests that subclone #33 contains cis-elements or closely linked genes that repress awn elongation. Bar length is 1 cm. (F) Frequency of awned seeds per panicle, (G) Awn length of transgenic lines of BAC sub-clone constructs. n.d.=not detected. #33 had shorter awns than #89. It is possible that cis-elements or genes which caused the repression of awn elongation are included in the longer segment of #33. Error bars represent standard deviation of the mean.

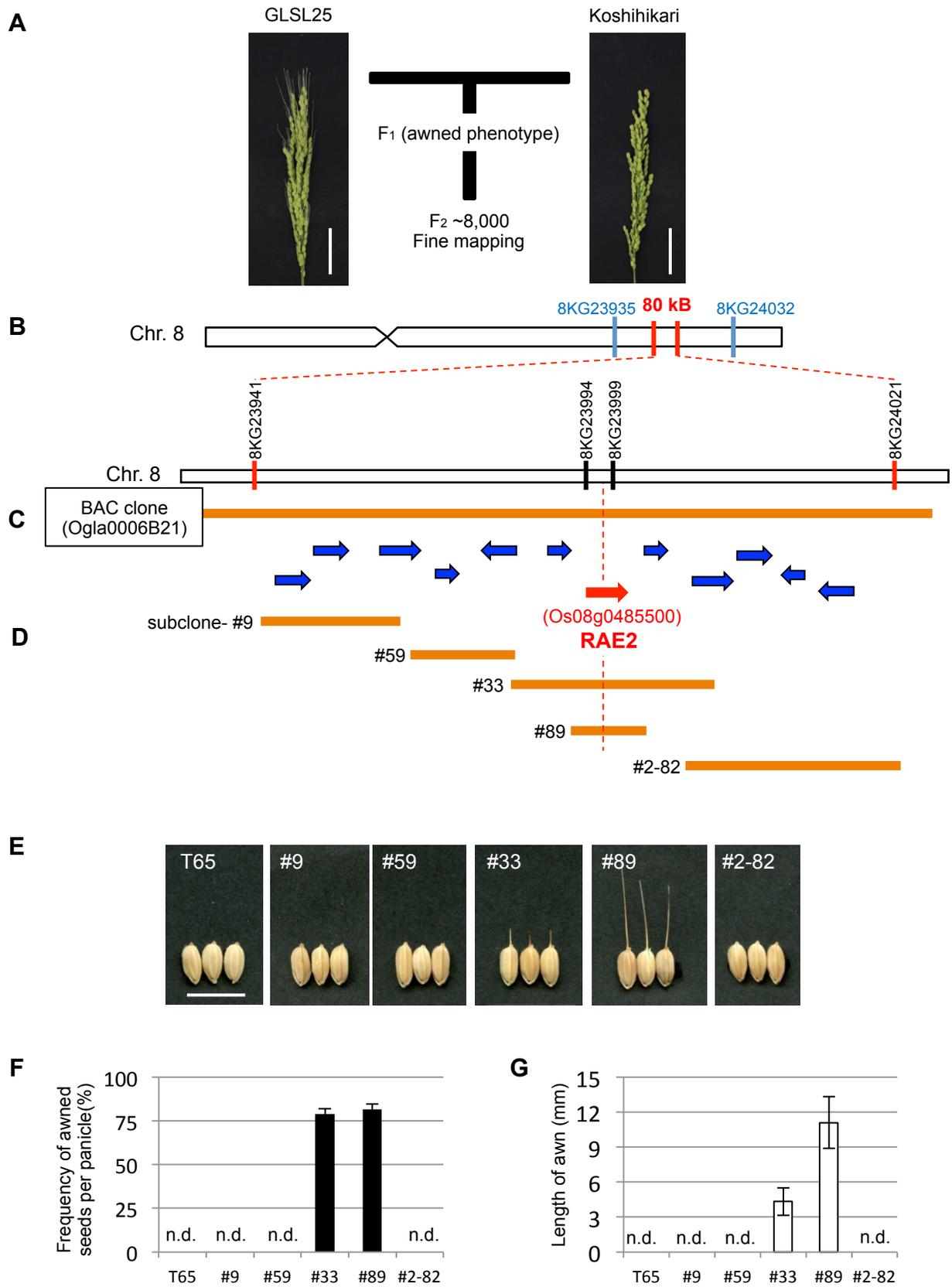


Fig. S1. Positional cloning and complementation test of *RAE2*.

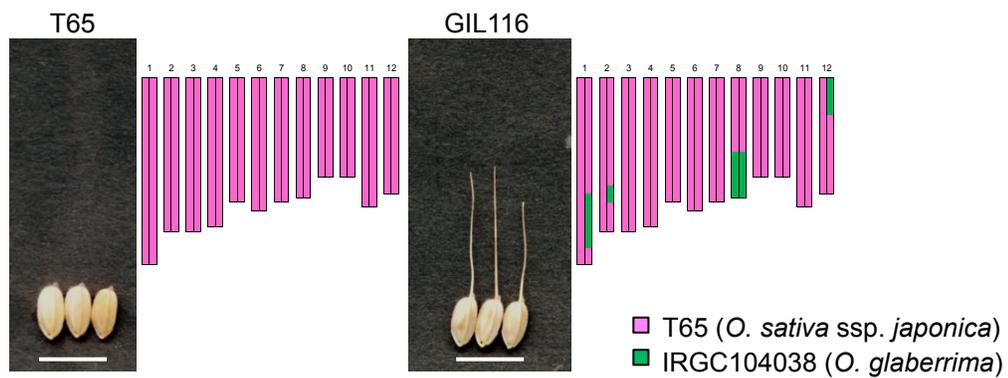


Fig. S2. Seed phenotype and graphical genotype of Taichung65 (T65) and the chromosome segment substitution line, GIL116.

Pink segment represents *O. sativa ssp. japonica* cv. Taichung65 (T65,) chromosomes and green segment represents *O. glaberrima* IRGC104038 chromosome segment. Bar length is 1 cm.

Fig. S3. Phylogenetic tree of EPF/EPFL family genes and comparison of the sequences in cysteine-rich region.

(A) Neighbor-joining phylogenetic tree of EPF/EPFL genes. Amino acid sequences for the cysteine-rich region were aligned using the ClustalW program. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1,000 replicates) is shown next to the branches (values 50% or greater are shown). The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method and are in the units of the number of amino acid substitutions per site. Evolutionary analyses were conducted in MEGA6. At: *Arabidopsis thaliana*, Os: *Oryza sativa*, Tu: *Triticum urartu*, Hv: *Hordeum vulgare*, Bd: *Brachypodium distachyon*. Each clade is consistent with previous report (24) except AtEPFL8 clade. Gray colored character indicates RAP-DB ID of *O. sativa* EPFLs. **(B)** Alignment of RAE2 cysteine-rich region peptide amino acid sequences with the half member of AtEPFL1-3 clade are shown in Fig. S3A. Pairs of cysteine (C) residues forming disulfide bonds predicted for *A. thaliana* EPF/EPFL genes are connected by lines. Two cysteine (C5 and C6) deletions in C-terminal region could be seen only in *rae2/OsEPFL1*.

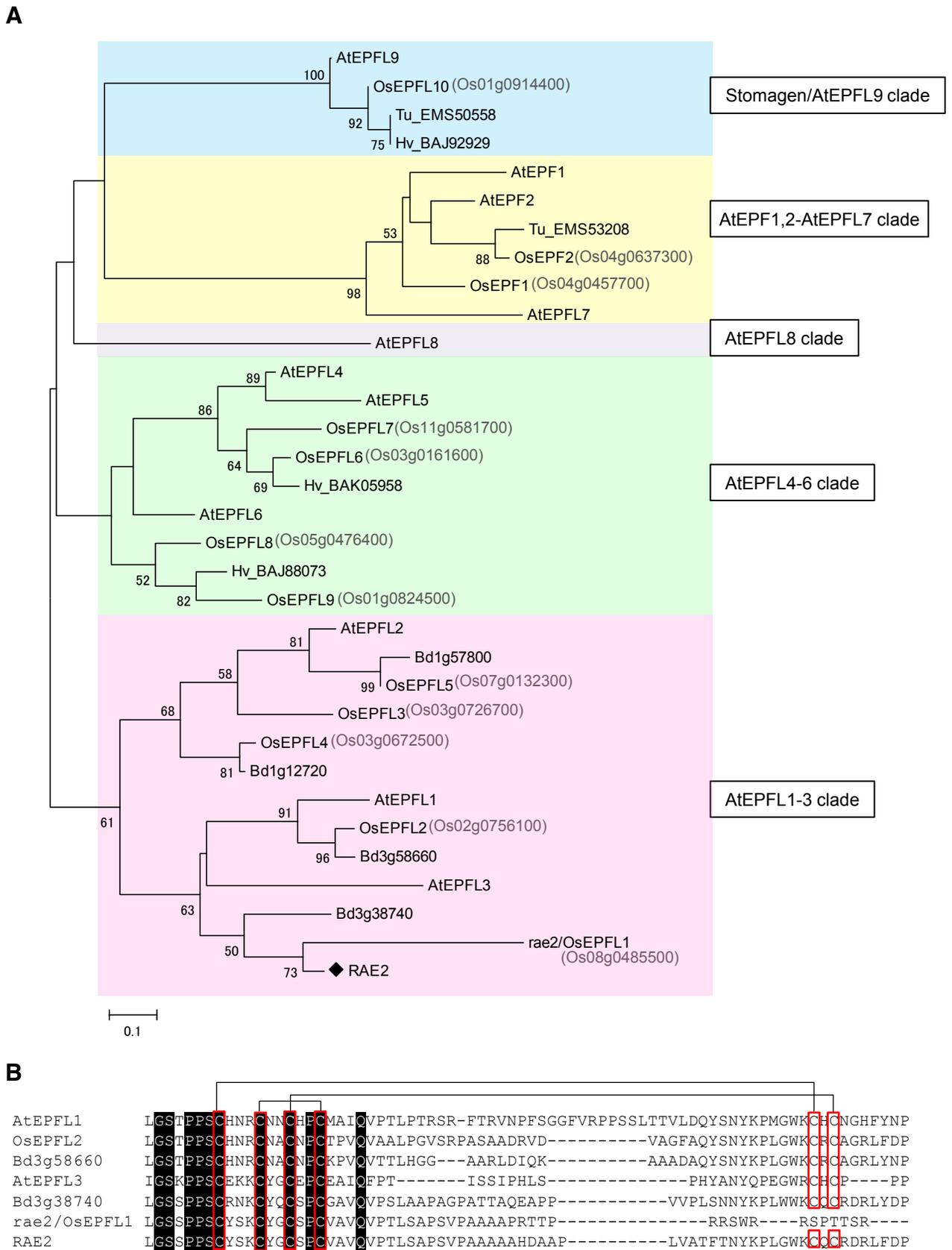


Fig. S3. Phylogenetic tree of EPF/EPFL family genes and comparison of the sequences in cysteine-rich region.

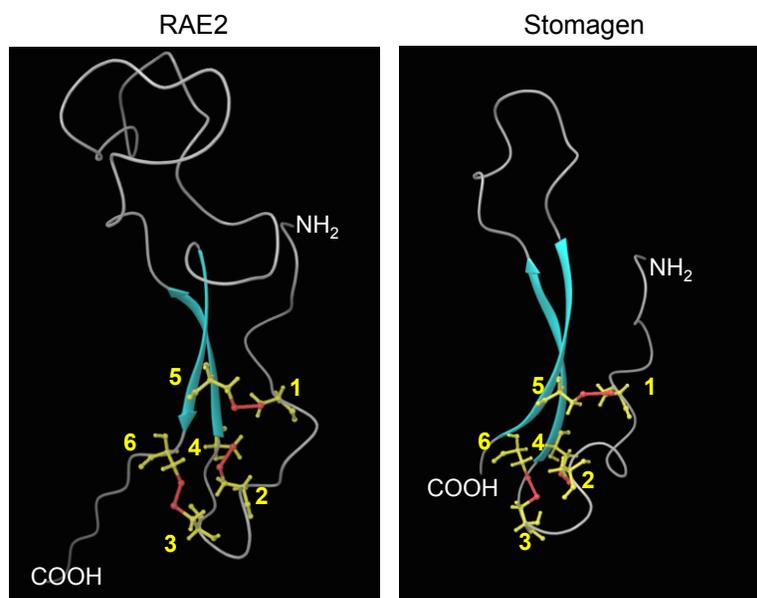


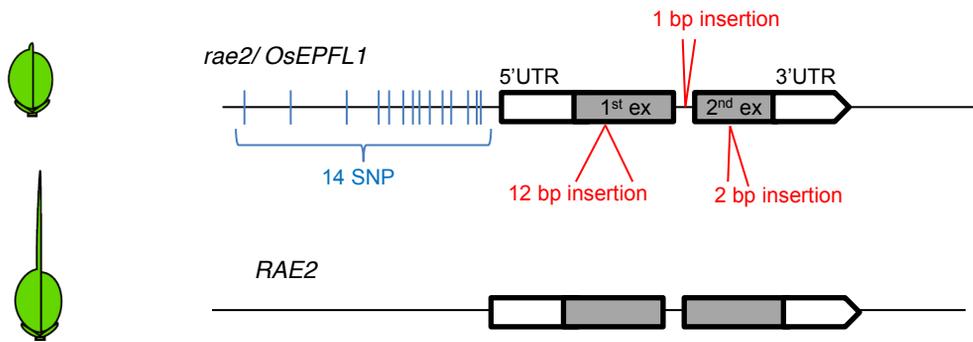
Fig. S4. Predicted 3-D structure of RAE2 and Stomagen.

Hypothetical 3-D structure of the mature peptide region of RAE2 and Stomagen modeled using the homology modeling system of the Mäestro (Schrödinger, NY, USA) software. Yellow: cysteine residues, red: disulfide bonds and atoms, blue: antiparallel beta sheets. Yellow number indicates the cysteine residue's location. The hypothetical structure of *rae2*/*OsEPFL1* could not be modelled because it lost two cysteine residues (C5 and C6) and did not fit the Stomagen template.

Fig. S5. RAE2 sequence comparison between *O. sativa* ssp. *japonica* cv. Koshihikari and *O. glaberrima* IRGC104038 .

(A) Schematic image of the *RAE2* and *rae2* gene. Gray-colored boxes in the gene model indicate exonic regions, white boxes indicate UTRs, and the line represents the promoter region, single intronic region and the terminator region of *RAE2*. Blue lines in the promoter region represent SNPs position and red lines represent insertion in *rae2/OsEPFL1* (Os08g048550) of Koshihikari (*O. sativa* ssp. *japonica*) compared with IRGC104038 (*O. glaberrima*). Koshihikari and Nipponbare have the same *rae2* sequence. (B) Comparison of the *RAE2* coding sequence in Koshihikari and IRGC104038. Koshihikari has 12 bp insertion in the first exon, and 2 bp insertion in the second exon as represented by the red square. The double red line represent GC-rich repeat region in Fig. S7A. (C) Comparison of the *RAE2* amino acid sequence between Koshihikari and IRGC104038. The insertion in the second exon caused a frameshift mutation in *rae2*.

A



B

rae2/OsEPFL1	1	ATGAGGACGGCGGCCACGCCGCTCTCGCCGCCCGCCCGCCGCGTTCGGCGCAGTGTTCCTCTCTGCGT	70
RAE2	1	ATGAGGACGGCGGCCACGCCGCTCTCGCCGCCCGCCCGCCGCGTTCGGCGCAGTGTTCCTCTCTGCGT	70
rae2/OsEPFL1	71	TGCTGCTCGCCTCGCCTCCGCCTCCGCCTCCAGGCTCCCTCCTCCTCGCCGTCTTCTTCCCTGGTTGG	140
RAE2	71	TGCTGCTCGCCTCGCCTCC-----AGGCTCCCTCCTCCTCGCCGTCTTCTTCCCTGGTTGG	128
rae2/OsEPFL1	141	TGGCGAGGTGGCGGTGGCGGTGGTGGCTGGGGAGGAGGAGAAGGTGCGGCTGGGGTCGAGCCCGCCGAGC	210
RAE2	129	TGGCGAGGTGGCGGTGGCGGTGGTGGCTGGGGAGGAGGAGAAGGTGCGGCTGGGGTCGAGCCCGCCGAGC	198
rae2/OsEPFL1	211	TGCTACAGCAAGTGCTACGGGTGCAGCCCGTGCCTCGCGGTGCAGGTGCCACCTTGTCGCCCCCGTCCG	280
RAE2	199	TGCTACAGCAAGTGCTACGGGTGCAGCCCGTGCCTCGCGGTGCAGGTGCCACCTTGTCGCCCCCGTCCG	268
rae2/OsEPFL1	281	TCCCGCCGCGCCCGCGCGCGGCACGACGCCGCGCCGCTCGTGGCGACGTTACCAACTACAAGCCGCTA	350
RAE2	269	TCCCGCCGCGCCCGCGCGCGGCACGACGCCGCGCCGCTCGTGGCGACGTTACCAACTACAAGCCGCTA	336
rae2/OsEPFL1	351	G-----	351
RAE2	337	GGTGGAAGTGCCAGTGCCGCGACCGCCTGTTTCGACCCCTGA	378

C

rae2/OsEPFL1	1	MRTAATPPLAAAAA VAVFLSALLLASASASASRLPPRRLLPLVGGEVAVAVVAGEEEKVRLGSSPPS	70
RAE2	1	MRTAATPPLAAAAA VAVFLSALLLASAS-----RLPPRRLLPLVGGEVAVAVVAGEEEKVRLGSSPPS	66
rae2/OsEPFL1	71	CYSKCYGCSPCAVQVPTLSAPSVAAAAAPRTTPRRSWRRSPTTSR	116
RAE2	67	CYSKCYGCSPCAVQVPTLSAPSVAAAAAHDAAPLVATFTNYKPLGWKCQCRDLFDP	125

Fig. S5. *RAE2* sequence comparison between *O. sativa* ssp. *japonica* cv. Koshihikari and *O. glaberrima* IRGC104038 .

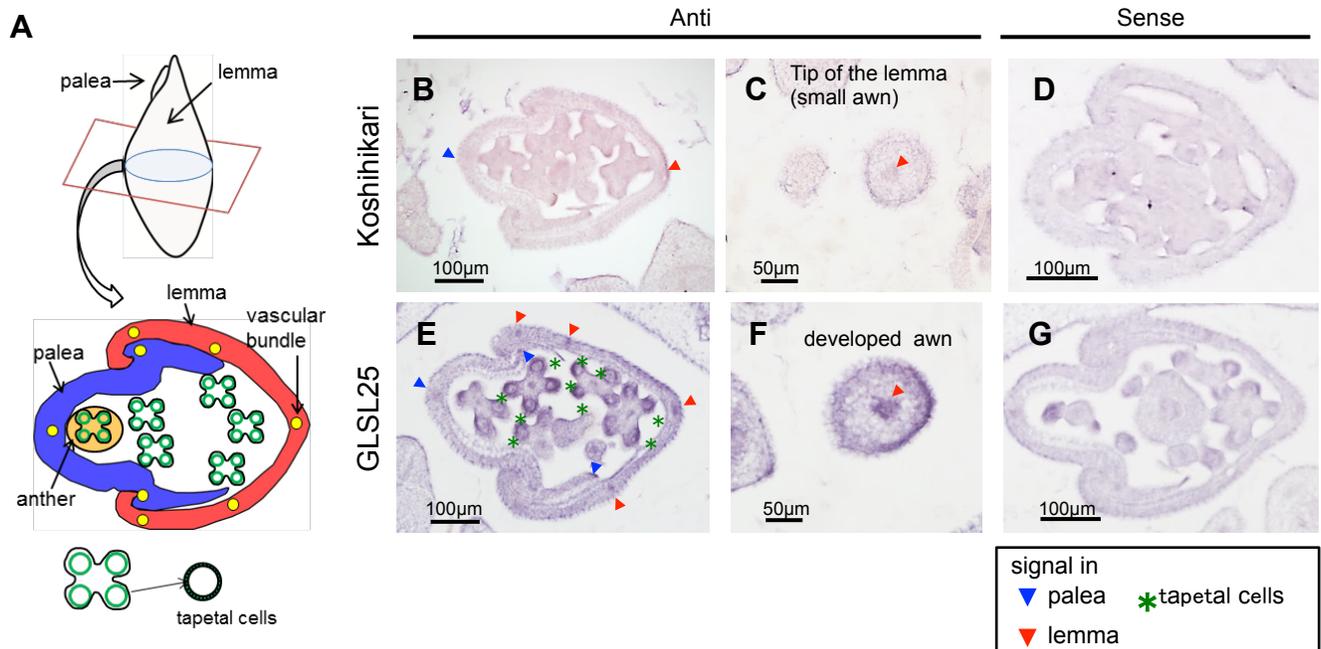


Fig. S6. Tissue-specific expression of RAE2 during awn development.

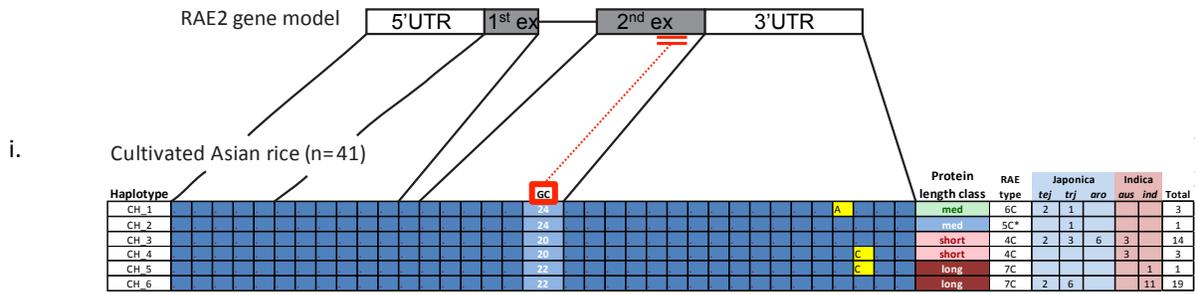
(A) Upper drawing shows a spikelet and lower drawing represents an image of the transverse section of a spikelet. (B-G) *in situ* hybridization with *rae2* antisense probe (B, C), *RAE2* probe (E, F) and sense probes (D, G) during the post Sp8 stage in Koshihikari and GLSL25. *RAE2* is expressed in the vascular bundle and outer layer of palea and lemma (indicated by blue and red arrow head, respectively). The strong signals are evident in the tapetal cells of stamens (indicated by green asterisk) (E). Expression of *RAE2* in the vascular bundle of Koshihikari is weaker compared to that of GLSL25 (C, F).

Fig. S7. RAE2 diversity and distribution across Asian and African rice accessions.

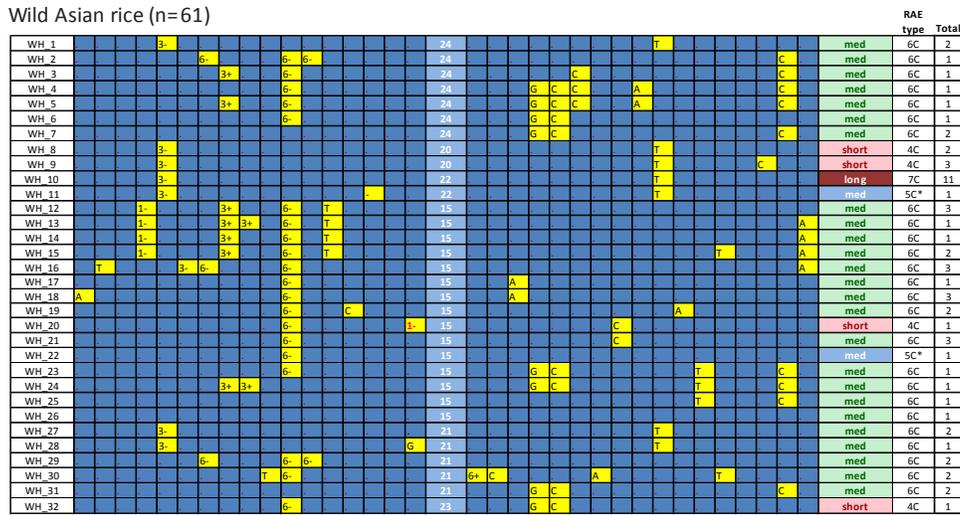
(A) RAE2 gene haplotypes across diverse Asian and African rice accessions. Polymorphic sites discovered within the coding and non-coding regions of RAE2 of 130 diverse rice accessions are shown in Table S2. (i) Cultivated Asian rice, (ii) Wild Asian rice, (iii) Wild and cultivated African rice.

Individual number in each population used for this analysis was noted in brackets (n=xx). Gray-colored boxes in the gene model shown at the top of the figure indicate exonic regions, white boxes indicate UTRs, and the line represents the single intronic region of RAE2. 'CH' indicates gene haplotypes present in cultivated Asian rice, 'WH' indicates gene haplotypes found in wild Asian rice while 'AFH' indicates gene haplotypes found in cultivated and wild African rice. 'RAE type' shows the number of cysteine residues predicted from translating the cDNA sequence. The value in column 'GC' represents the number of nucleotides within a highly variable GC-rich region of the RAE2 second exon (indicated in double red line as same as shown in Fig. S5B). At all other sites, blue cells represent alleles that match the reference genome (cv. Nipponbare) while yellow cells are non-reference alleles. The numbers of accessions that harbored each haplotype are indicated in the right-hand table (tej = temperate japonica, trj = tropical japonica, aro = aromatic, ind = indica, aus = aus, determined by HDRA genome-wide information). **(B)** Distribution of translated propeptide lengths across 107 Asian rice accessions and 23 African rice accessions. We defined the protein length class depending on amino acid length; short=110 ~ 120 AA, medium=120~135 AA, long=195~200 AA (red=4C/short, blue=5C/medium, green=6C/medium, yellow=7C/long).

A



ii.



iii.



B

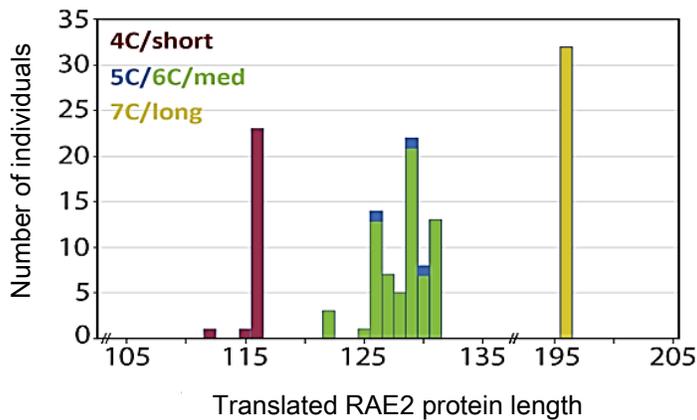


Fig. S7. RAE2 diversity and distribution across Asian and African rice accessions.

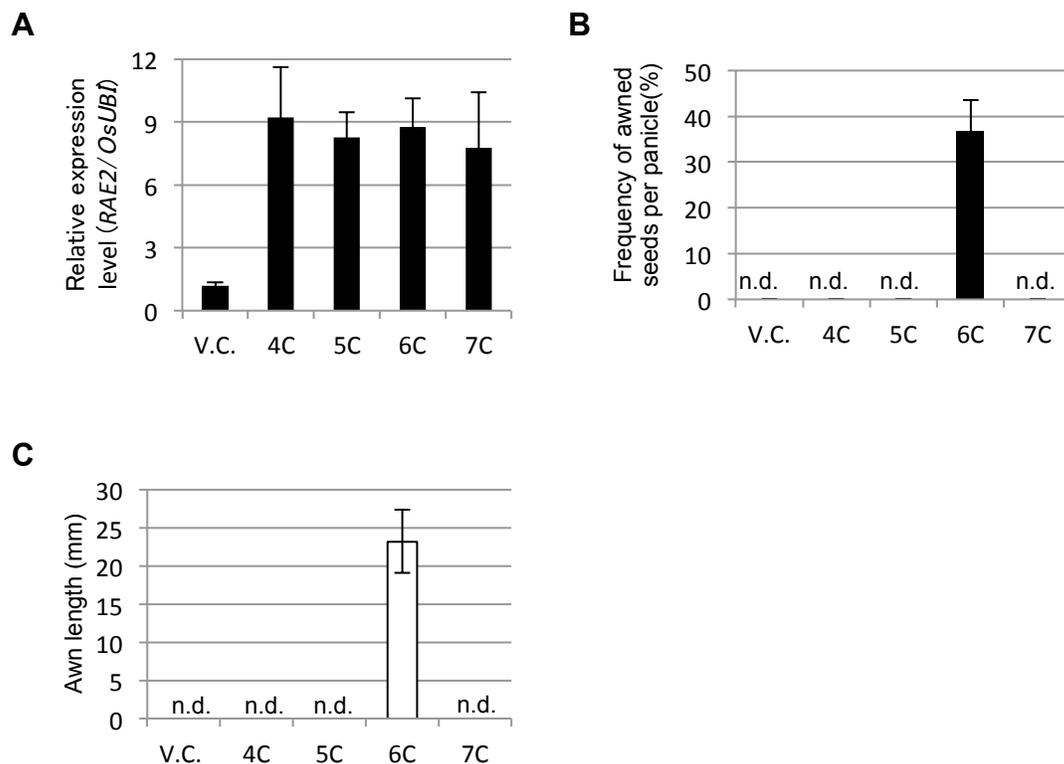


Fig. S8. Different types of *RAE2* and definition of each function for awn elongation.

(A) Relative expression levels of *RAE2* in young panicles of transgenic lines of overexpression construct; 4C, 5C, 6C, 7C. *OsUBI* used as internal control. (B) Frequency of awned seeds per panicle, (C) awn length in each overexpression lines. We measured 10 transgenic plants per each line. n.d.=not detected. Error bars represent standard deviation of the mean. We used NSFTV223 (listed in Table S4) genomic DNA for making 5C construction in this experiments. Since our study suggested that conserved cysteine residues number is important for *RAE2* conformation, we expect the other two singletons of 5C type *RAE2* (derived from NSFTV673 and NSFTV762) lose those function.

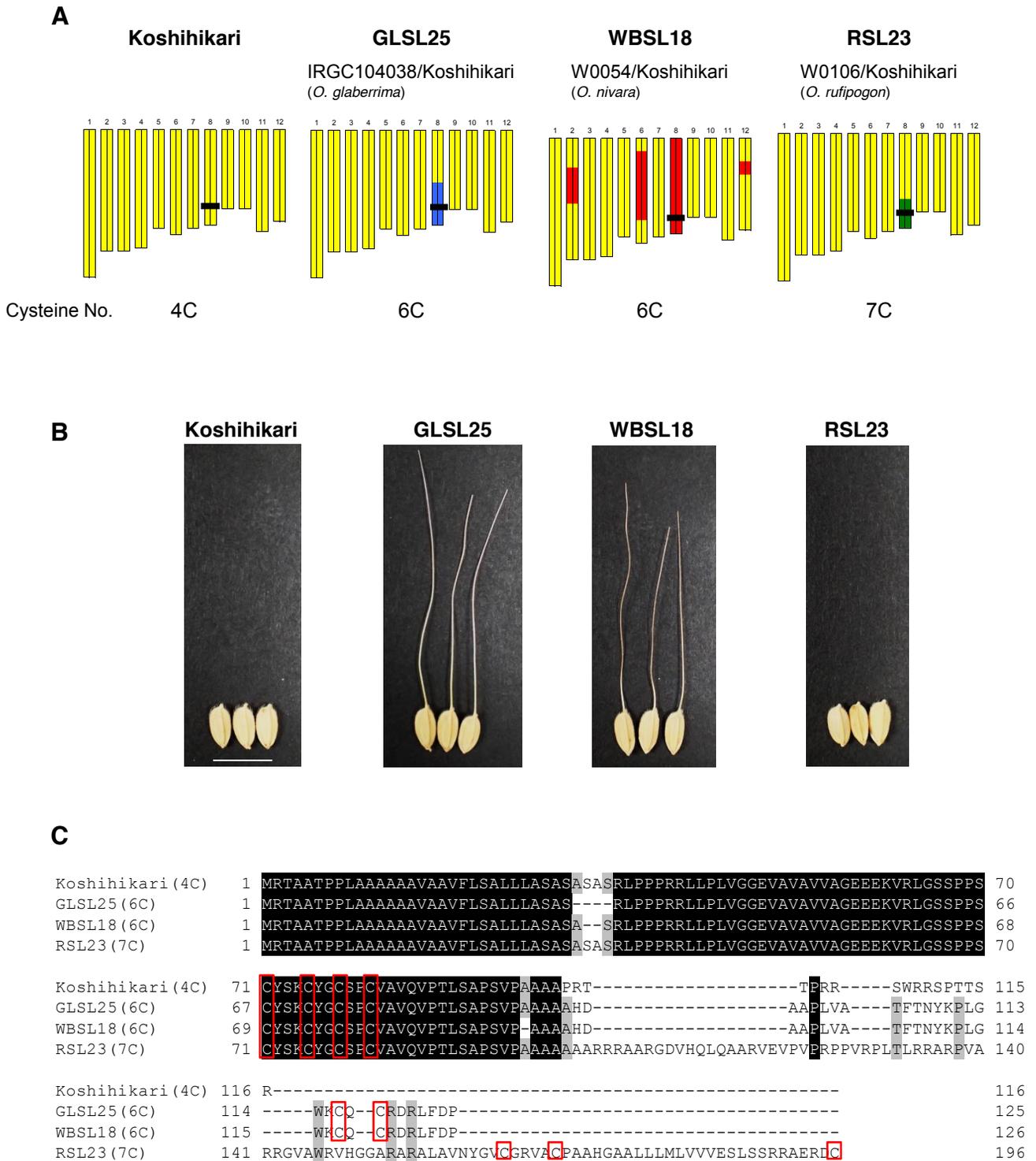


Fig. S9. Awn phenotypes of CSSLs and the number of RAE2 cysteine residues.

Fig. S10. The RAE2 maturation process occurs specifically in the spikelet.

(A) The model of secretory peptide cleavage. EPF/EPFL family peptides are composed of a signal peptide region (sp, blue), pro-peptide region (pro, green) and mature peptide region (ma, pink). After transcription, the whole peptide called pre-pro-peptide, is cleaved at the border of the signal peptide sequence and then forms the pro-peptide. Disulfide bond formation and folding occurs in the endoplasmic reticulum before the pro-peptide is secreted into the extracellular matrix. Further cleavage occurs in the border between the pro-peptide and mature peptide region.

(B) The membrane stained by ponceau S (Wako, Japan) as loading control of Fig.4A. Total protein amount is 30 μ g in each lane.

(C) The procedure of *in vitro* processing assay. Scissors drawing is the image of protease cleaving the RAE2 cleavage site.

(D) *in vitro* processing assay of recombinant RAE2 pro-peptide fused with 3xFLAG tag incubated with plant extracts of Koshihikari spikelet or buffer (mock). Fusion peptide and cleaved peptides were detected by anti-RAE2 antibody. The other details are same as described in Fig. 4C.

(E) *in vitro* processing assay with or without protease inhibitor cocktail, Complete (Roche, Basle). SE= spikelet extract. The other description is same as described.

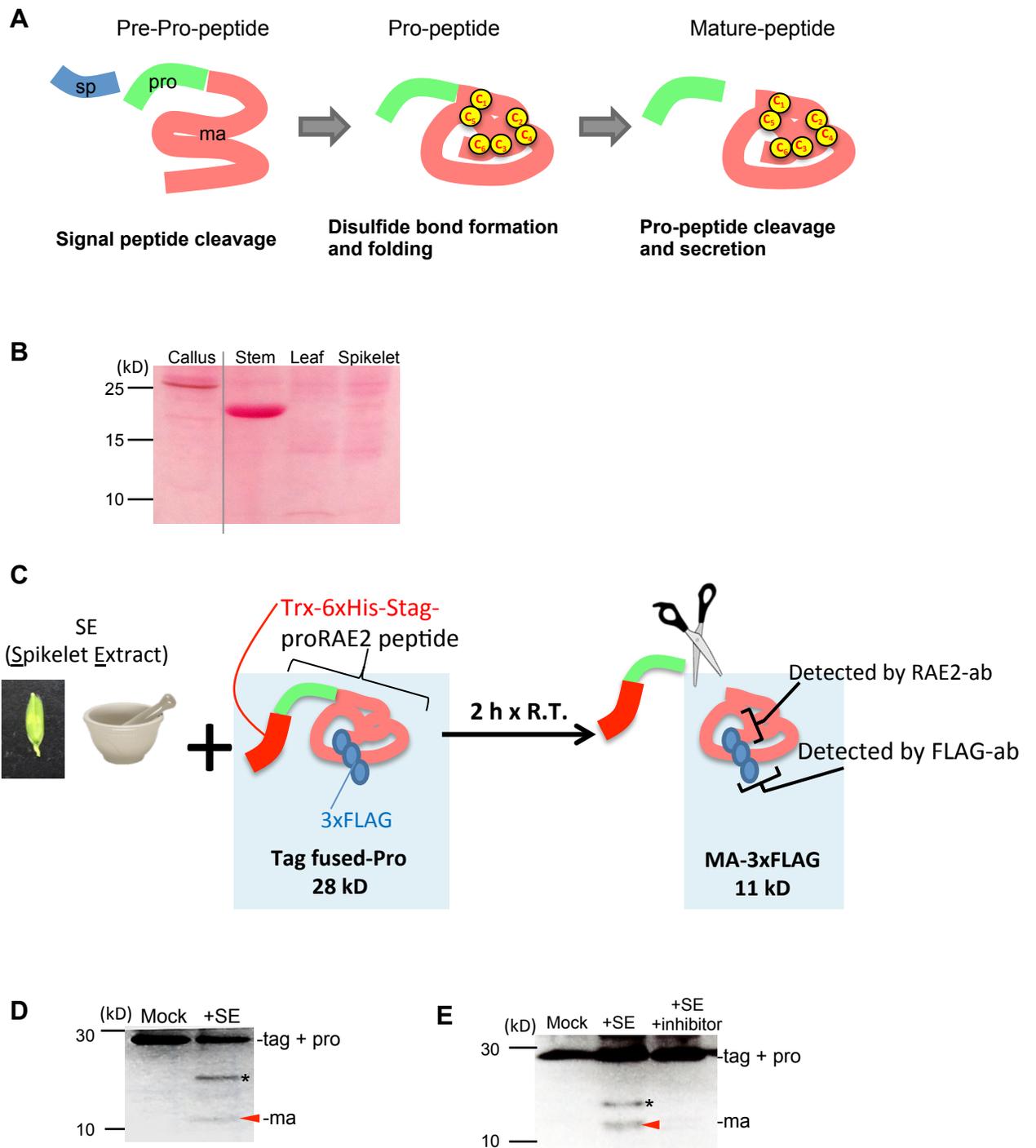


Fig. S10. The RAE2 maturation process occurs specifically in the spikelet.

Fig. S11. Specific expression pattern of SLP1.

(A) Spatio-temporal expression data of *SLP1* from the rice expression database, RiceXpro (<http://ricexpro.dna.affrc.go.jp/>). Samples were used for hybridization using the Agilent one-color (Cy3) microarray-based gene analysis system. The expression profile is shown as raw data representing Cy3 signal intensity and normalized data (log2). *SLP1* was specifically expressed in the young inflorescence (0.6 mm – 1.0 mm, green colored bar) and moderately expressed in the lemma and palea (4.0 mm – 7.0 mm, blue colored bar). (B) semi-quantitative PCR result of *SLP1* by primers KU144 and KU146 (Table S7). The template which is extracted from Koshihikari is the same as the one used in qRT-PCR in Fig. 2A. LS= leaf sheath, LB= leaf blade, IN= internode, RO= root, PA= young panicle. We used *OsACTIN1* as internal control. (C) The ion spectrum of SLP1 peptide: 543-SAIMTTAITGDNDKIR-560 identified after iTRAQ labeling and LC-MALDI MS/MS analysis using ProteinPilot software 5.0 AB SCIEX. (i) Identified peptide using detected b- and y-- ion series, (ii) MS/MS fragmentation spectra, (iii) Quantitation of iTRAQ report ions: The peak of #1 and #2 were extracted from Koshihikari spikelets labeled with 114, 115 iTRAQ reagents respectively. Tryptic peptides from *in vitro*-translation SLP1, and synthetic SLP1 peptide were labeled with 116, 117 iTRAQ reagents respectively.

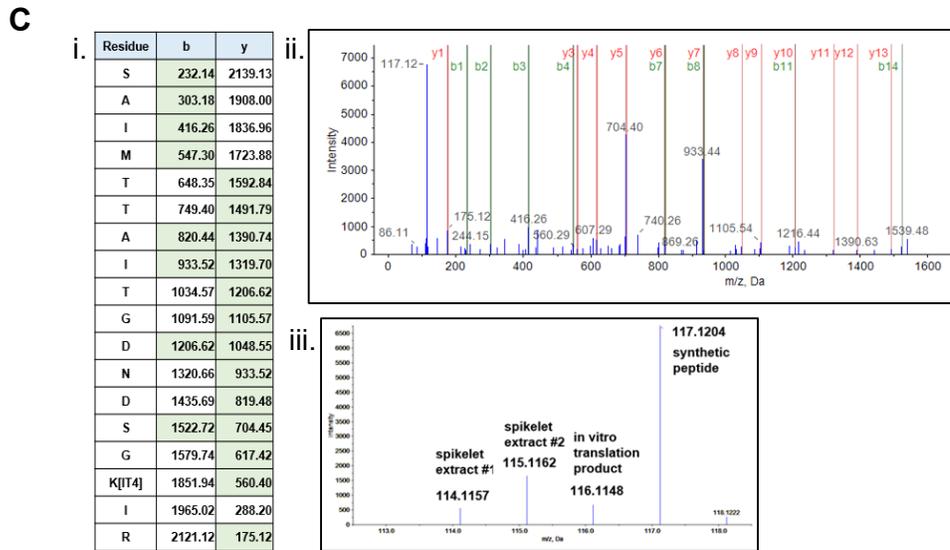
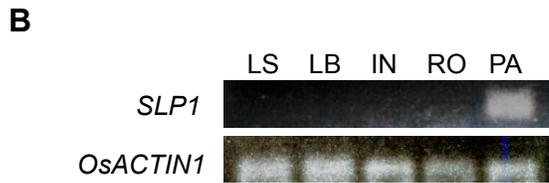
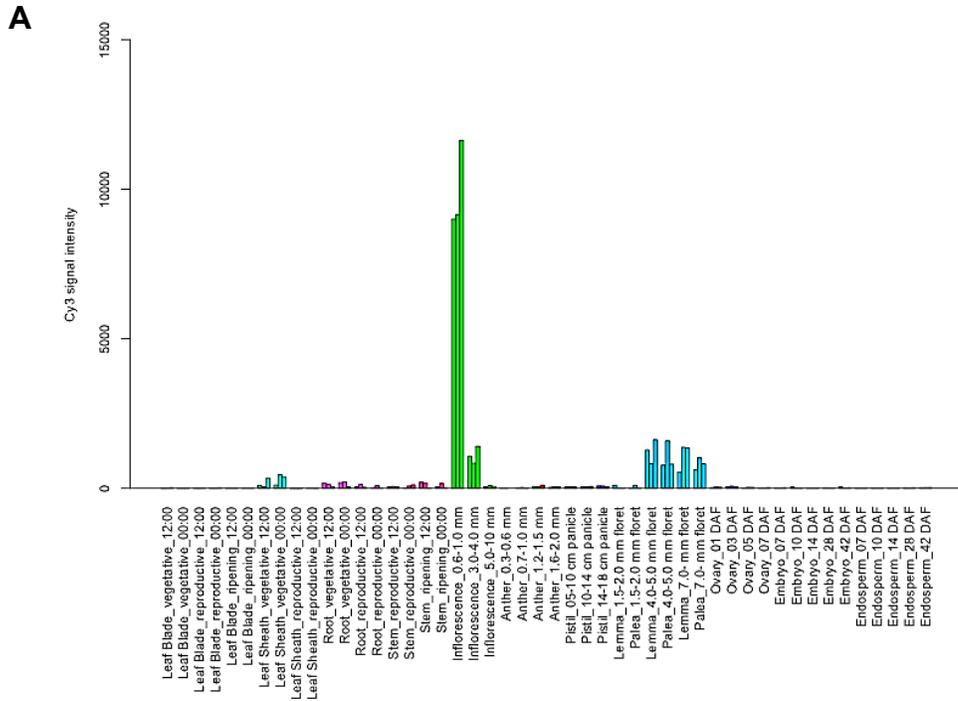
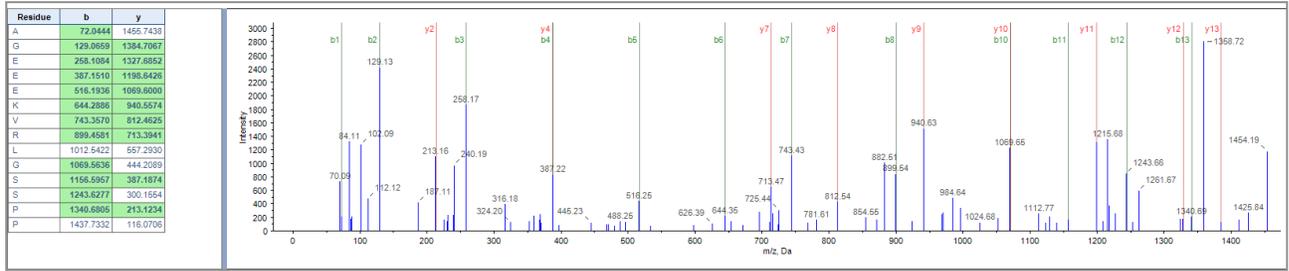


Fig. S11. Specific expression pattern of SLP1 .

Fig. S12. Detection of the cleavage site of RAE2 by *in vitro* processing assay.

(A) The product ion spectrum for the synthetic RAE2 peptide after *in vitro* processing reactions incubated with SLP1 synthesized by *in vitro* translation (correlated with Fig.4D). The production of spectra are annotated for b and y ion series, using the Paragon algorithm and the left table shows the identification results for the peptides using ProteinPilot 5.0. (B) Mass spectrometry of *in vitro* processing reactions of the series of amino acid substituted synthetic peptides of RAE2 incubated with *in vitro*-synthesized SLP1. Left lane (entitled cleaved) and right lane (entitled uncleaved) showed that mu-synRAE2 which could or could not be cleaved by SLP1 respectively. Red tangle indicates the fragment cleaved between P65 and S66. The table in lower left shows the summary of the result; o and x are consistent with cleaved and uncleaved. There are two peaks which we infer the reason described in Fig.4D (* = 54-EEEKVRLGSSPPSCYSK-70, ** = 54-EEEKVRLGSSPP-65). Incubation of panicle extracts and synthetic RAE2 peptides represented same results using SLP1 product of *in vitro* translation (data not shown). SLP1 was unable to cleave synthetic peptides with mutations around the amino acid positions P65 and S66

A



B

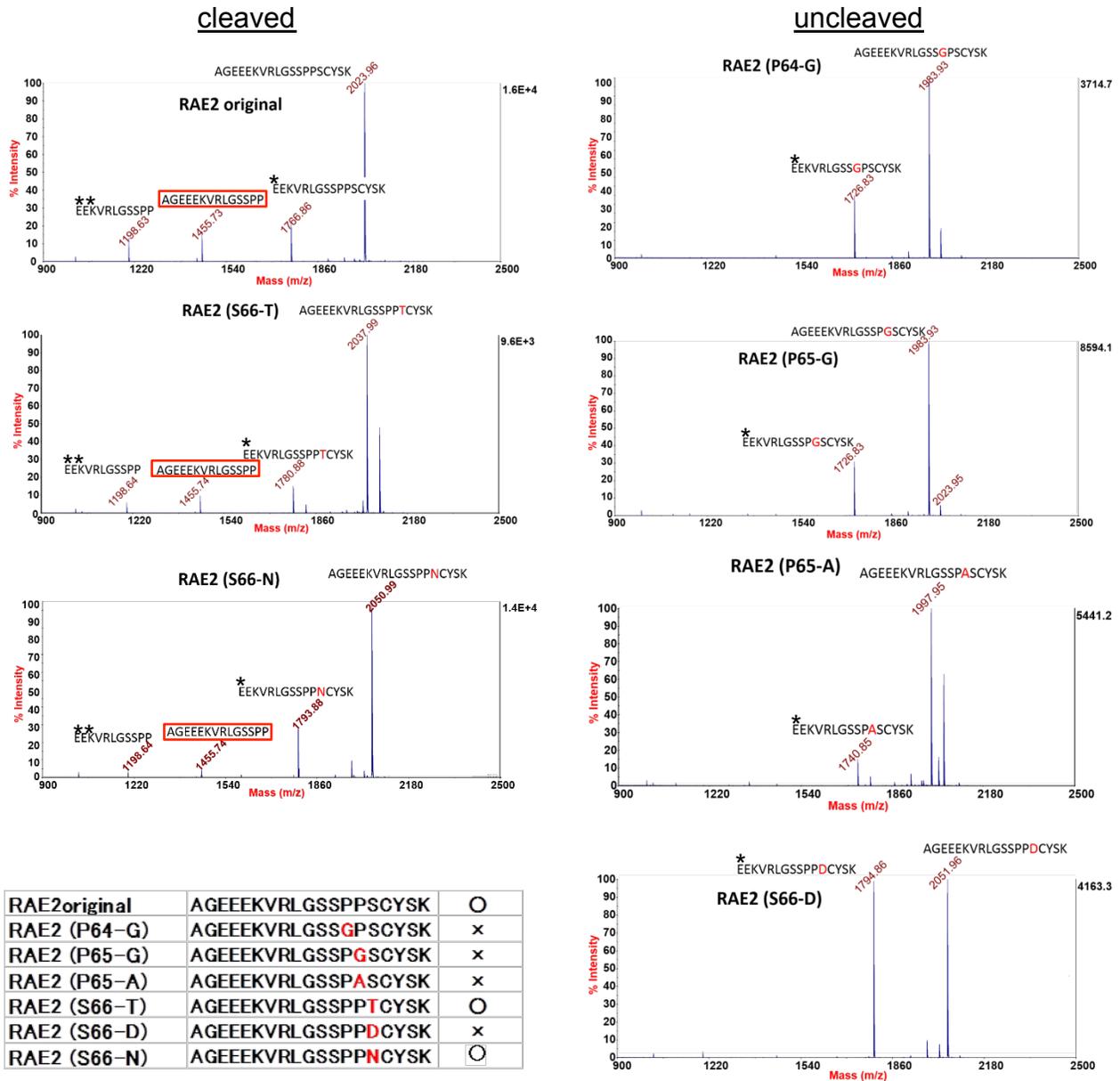
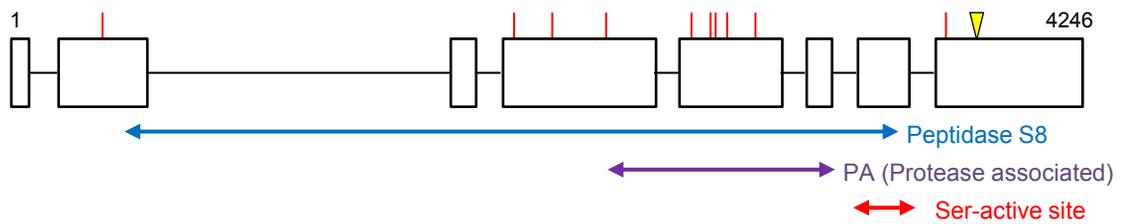


Fig. S12. Detection of the cleavage site of RAE2 by *in vitro* processing assay.

Fig. S13. Sequence and amino acid structure of SLP1.

(A) Schematic image of SLP1. White-colored boxes in the gene model indicate exonic regions, and the line represents the intronic regions of SLP1. Red bars represent SNPs position and yellow triangle represent insertion in CG14 (*O. glaberrima*) compared with Nipponbare (*O. sativa* ssp. *japonica*). Blue arrow represents the peptidase S8 domain, purple arrow represents the protease-associated domain, red arrow represents the serine-active site according to Swiss-prot. (B) Sequence comparison of SLP1 amino acid in Nipponbare (*O. sativa* ssp. *japonica*) and CG14 (*O. glaberrima*). Blue square shows the peptidase S8 domain, the purple and red arrows represent the protease-associated domain and the serine-active site, respectively.

A



B

O. sativa	1	MQTYVIVFDGLPASPSGLLATVVTFSQLLYVLSPIQVIVVQIDESFVGVKQLPGVLAVIPDVLHKVHTT	70
O. glaberrima	1	MQTYVIVFDGLPASPSGLLATVVTFSQLLYVLSPIQVIVVQIDESFVGVKQLPGVLAVIPDVLHKVHTT	70
O. sativa	71	RSWDFLELERNGAATGAWKDAAKYG/DAIIGNVDTGVWPESASFKDDGYSVPSRWRGKCITGNDTTFKCN	140
O. glaberrima	71	RSWDFLELERNGAATGAWKDAAKYG/DAIIGNVDTGVWPESASFKDDGYSVPSRWRGKCITGNDTTFKCN	140
O. sativa	141	NKLI GAGFFNLGFLASGLLQ GKPPSQAAELYTPRDYIGHGHTLSTAGGGFVPDASVFGHGKTAKGGSP	210
O. glaberrima	141	NKLI GAGFFNLGFLASGLLQ GKPPSQAAELYTPRDYIGHGHTLSTAGGGFVPDASVFGHGKTAKGGSP	210
O. sativa	211	LARVAAYKACYAEGCSSSDILAAMVTAVEDG VNVLSLSVGGPADDYLSDP IAI GAFYAVQKGVIVVCSAS	280
O. glaberrima	211	LARVAAYKACYAEGCSSSDILAAMVTAVEDG VNVLSLSVGGPADDYLSDP IAI GAFYAVQKGVIVVCSAS	280
O. sativa	281	NSGPQPGSVTNVAPWILTVGASTMDRDFPAYVTFGGVTSSMTIKGQSLSNSTLPQGQRYAMINAKNANAA	350
O. glaberrima	281	NSGPQPGSVTNVAPWILTVGASTMDRDFPAYVTFGGVTSSMTIKGQSLSNSTLPQGQRYAMINAKNANAA	350
O. sativa	351	NVPSENSTLCFPGSLDSKVRGKIVVCTRGVNARVEKGLVVKQAGGVGMVLCN YAGNGEDVIADPHLIAA	420
O. glaberrima	351	NVPSENSTLCFPGSLDSKVRGKIVVCTRGVNARVEKGLVVKQAGGVGMVLCN YAGNGEDVIADPHLIAA	420
O. sativa	421	AHVSYSQCINLFN YLGSTDNPVGYITASDARLGVKPAPVMAAFSSRGPNPITPQILKPDITAPGVSVIAA	490
O. glaberrima	421	AHVSYSQCINLFN YLGSTDNPVGYITASDARLGVKPAPVMAAFSSRGPNPITPQILKPDITAPGVSVIAA	490
O. sativa	491	YSEAVSPELSFDDRRVPYNIMSGTSMSCPHVSGIVGLIKTKYPDWT PAMIKSAIMTTAITGDNDSGKIR	560
O. glaberrima	491	YSEAVSPELSFDDRRVPYNIMSGTSMSCPHVSGIVGLIKTKYPDWT PAMIKSAIMTTAITGDNDSGKIR	560
O. sativa	561	DETGAAATPFA YGSGHVRSVQALDPGLVYD TTSADYADFLCALRPTQN --PLPLPVFGDDGKPRACSQGA	628
O. glaberrima	561	DETGAAATPFA YGSGHVRSVQALDPGLVYD TTSADYADFLCALRPTQN --PLPLPVFGDDGKPRACSQGA	630
O. sativa	629	QYGRPEDLNYP SIAVPCLSGSATVRRRVKNVGAAPCRYAVSVTEALAGVKVTVPPELSFESYGEEREFT	698
O. glaberrima	631	QYGRPEDLNYP SIAVPCLSGSATVRRRVKNVGAAPCRYAVSVTEALAGVKVTVPPELSFESYGEEREFT	700
O. sativa	699	VRLEVQDAAAAANYVFGSIEWSEESESDPDRKHRVRSPIVAKTTCG	744
O. glaberrima	701	VRLEVQDAAAAANYVFGSIEWSEESESDPDRKHRVRSPIVAKTTCG	746

Fig. S13. Sequence and amino acid structure of SLP1.

Table S1. The awn phenotype of transgenic lines and parental lines.

	Number of seeds		awn length(mm)	Frequency of awned seeds(%)
	Awned	Awnless		
WCSL26	237	14	26.6±0.84*	96.1±6.52*
Nipponbare	0	261	0	0
pGW501(V.C.) (line1)	0	223	0	0
pGW501(V.C.) (line2)	0	129	0	0
pGW501(V.C.) (line3)	0	217	0	0
pRAE2::RAE2 (line 1)	230	11	25.2±0.73*	95.4±3.43*
pRAE2::RAE2 (line 2)	259	36	28.3±0.62*	87.8±4.25*
pRAE2::RAE2 (line 3)	213	12	15.1±0.68*	94.7±8.15*
GIL116	143	143	20.2±2.82	100±0.0
T65	0	242	0.0±0.0	0.0±0.0
pANDA(V.C) (line1)	49	1	16.8±3.09	98.3±1.08
pANDA(V.C) (line2)	55	0	17.5±5.38	100±0.0
pANDA(V.C) (line3)	12	0	23.2±3.56	100±0.0
OgAWN2-RNAi (line 1)	25	5	10.1±2.06*	83.3±8.15
OgAWN2-RNAi (line 2)	23	3	7.4±3.04*	84.5±6.53
OgAWN2-RNAi (line 3)	26	6	8.5±2.13*	81.3±4.78
pCAMBIA(V.C.) (line1)	0	41	0	0
pCAMBIA(V.C.) (line2)	0	35	0	0
pCAMBIA(V.C.) (line3)	0	33	0	0
rae2(4C)ox (line1)	0	38	0	0
rae2(4C)ox (line2)	0	42	0	0
rae2(4C)ox (line3)	0	45	0	0
rae2(5C)ox (line1)	0	50	0	0
rae2(5C)ox (line2)	0	32	0	0
rae2(5C)ox (line3)	0	34	0	0
RAE2(6C)ox (line1)	25	15	23.8±5.24*	38.3±1.15*
RAE2(6C)ox (line2)	29	18	19.8±2.56*	40.2±3.43*
RAE2(6C)ox (line3)	24	13	24.5±3.17*	37.5±2.89*
rae2(7C)ox (line1)	0	48	0	0
rae2(7C)ox (line2)	0	51	0	0
rae2(7C)ox (line3)	0	39	0	0

* P < 0.05 based on two-tailed Student's t test.

Table S2. Germplasm information used for *RAE2* diversity study.

Accession ID	Name	Species: subpopulation	Germplasm Repository	Origin	Awn Class (SES)	GC length	RAE2 protein length class
NSFTV7	Arias	<i>O. sativa: tropical japonica</i>	GSOR 301007	Indonesia	1	22	long/7C
NSFTV30	Chiem Chanh	<i>O. sativa: indica</i>	GSOR 301028	Vietnam	1	22	long/7C
NSFTV46	Dourado Agulha	<i>O. sativa: tropical japonica</i>	GSOR 301043	Brazil	5	20	short/4C
NSFTV50	DZ78	<i>O. sativa: aus</i>	GSOR 301046	Bangladesh	0	20	short/4C
NSFTV53	Firooz	<i>O. sativa: aromatic</i>	GSOR 301049	Iran	0	20	short/4C
NSFTV59	Gogo Lempuk	<i>O. sativa: tropical japonica</i>	GSOR 301055	Indonesia	9	24	med/6C
NSFTV75	Jambu	<i>O. sativa: tropical japonica</i>	GSOR 301068	Indonesia	5	20	short/4C
NSFTV93	Kitrana 508	<i>O. sativa: aromatic</i>	GSOR 301085	Madagascar	9	20	short/4C
NSFTV105	Mehr	<i>O. sativa: aus</i>	GSOR 301097	Iran	0	20	short/4C
NSFTV112	N12	<i>O. sativa: aromatic</i>	GSOR 301104	India	5	20	short/4C
NSFTV142	Shai-Kuh	<i>O. sativa: indica</i>	GSOR 301133	China	7	22	long/7C
NSFTV153	T26	<i>O. sativa: aus</i>	GSOR 301144	India	7	20	short/4C
NSFTV154	Ta Hung Ku	<i>O. sativa: temperate japonica</i>	GSOR 301145	China	9	24	med/6C
NSFTV160	NSF-TV 160	<i>O. sativa: aromatic</i>	GSOR 301151	Iran	1	20	short/4C
NSFTV173	Nipponbare	<i>O. sativa: temperate japonica</i>	GSOR 301164	Japan	1	20	short/4C
NSFTV200	P 737	<i>O. sativa: aus</i>	GSOR 301191	Pakistan	3	20	short/4C
NSFTV221	Sadri Belyi	<i>O. sativa: aromatic</i>	GSOR 301212	Azerbaijan	0	20	short/4C
NSFTV222	Paraiba Chines Nova	<i>O. sativa: indica</i>	GSOR 301213	Brazil	0	22	long/7C
NSFTV223	Priano Guaira	<i>O. sativa: tropical japonica</i>	GSOR 301214	Brazil	1	24	med/5C
NSFTV243	Tropical Rice	<i>O. sativa: temperate japonica</i>	GSOR 301233	Ecuador	0	22	long/7C
NSFTV250	Bulgare	<i>O. sativa: temperate japonica</i>	GSOR 301240	France	9	24	med/6C
NSFTV261	Shim Balte	<i>O. sativa: aus</i>	GSOR 301251	Iraq	9	20	short/4C
NSFTV265	Vialone	<i>O. sativa: temperate japonica</i>	GSOR 301255	Italy	3	22	long/7C
NSFTV269	Sundensis	<i>O. sativa: indica</i>	GSOR 301259	Kazakhstan	1	22	long/7C
NSFTV284	IR-44595	<i>O. sativa: indica</i>	GSOR 301274	Nepal	1	22	long/7C
NSFTV298	LD 24	<i>O. sativa: indica</i>	GSOR 301288	Sri Lanka	0	22	long/7C
NSFTV309	Manzano	<i>O. sativa: tropical japonica</i>	GSOR 301299	Zaire	0	20	short/4C
NSFTV310	R 101	<i>O. sativa: tropical japonica</i>	GSOR 301300	Zaire	0	22	long/7C
NSFTV337	Sabharaj	<i>O. sativa: indica</i>	GSOR 301327	Bangladesh	0	22	long/7C
NSFTV339	Yodanya	<i>O. sativa: indica</i>	GSOR 301329	Myanmar	0	22	long/7C
NSFTV349	Chang Ch'Sang Hsu Tao	<i>O. sativa: indica</i>	GSOR 301339	China	0	22	long/7C
NSFTV356	JC 117	<i>O. sativa: indica</i>	GSOR 301344	India	0	22	long/7C
NSFTV369	Sathi	<i>O. sativa: aus</i>	GSOR 301356	Pakistan	1	20	short/4C
NSFTV373	Lambayeque 1	<i>O. sativa: aromatic</i>	GSOR 301360	Peru	0	20	short/4C
NSFTV377	PR 304	<i>O. sativa: tropical japonica</i>	GSOR 301362	Puerto Rico	0	22	long/7C
NSFTV379	Wanica	<i>O. sativa: tropical japonica</i>	GSOR 301364	Suriname	0	22	long/7C
NSFTV380	Tainan-lku No. 512	<i>O. sativa: temperate japonica</i>	GSOR 301365	Taiwan	0	20	short/4C
NSFTV381	325	<i>O. sativa: tropical japonica</i>	GSOR 301366	Taiwan	9	22	long/7C
NSFTV395	OS 6 (WC 10296)	<i>O. sativa: tropical japonica</i>	GSOR 301378	Zaire	1	22	long/7C
NSFTV398	93-11	<i>O. sativa: indica</i>	GSOR 301399	China	3	22	long/7C
NSFTV399	Spring	<i>O. sativa: tropical japonica</i>	GSOR 301381	United States	1	20	short/4C
NSFTV400	Yang Dao 6	<i>O. sativa: indica</i>	GSOR 301400	China	7	22	long/7C
NSFTV402		<i>O. spontanea</i>	IRGC80539	India	9	20	short/4C
NSFTV410		<i>O. nivara</i>	IRGC80759	Myanmar	9	21	med/6C
NSFTV413		<i>O. nivara</i>	IRGC81850	India	9	15	med/6C
NSFTV415		<i>O. spontanea</i>	IRGC81909	India	9	15	med/6C
NSFTV416		<i>O. spontanea</i>	IRGC81970	Thailand	0	22	long/7C
NSFTV422		<i>O. rufipogon</i>		Vietnam		15	med/6C
NSFTV427		<i>O. rufipogon</i>		China	9	21	med/6C
NSFTV431		<i>O. rufipogon</i>	IRGC82992	China	9	15	med/6C
NSFTV432		<i>O. rufipogon</i>		Thailand		21	long/7C
NSFTV433		<i>O. rufipogon</i>	IRGC83795	India	9	23	short/4C
NSFTV435		<i>O. rufipogon</i>	IRGC86448	Thailand		24	med/6C
NSFTV438 (438_B2_1_S2)		<i>O. rufipogon</i>		India	9	22	long/7C
NSFTV443		<i>O. nivara</i>	IRGC93183	Nepal	9	15	med/6C
NSFTV444		<i>O. nivara</i>	IRGC93188	Nepal	9	15	med/6C
NSFTV446		<i>O. spontanea</i>	IRGC93224	Nepal	9	15	med/6C
NSFTV450		<i>O. nivara</i>	IRGC100916	China	9	15	med/6C
NSFTV453		<i>O. rufipogon</i>	IRGC103404	Bangladesh	9	22	long/7C
NSFTV457 (457_B3_1_S2)		<i>O. nivara</i>		Bangladesh	9	27	med/6C
NSFTV461 (461_A1_1_S2)		<i>O. rufipogon</i>		China	9	20	short/4C
NSFTV467		<i>O. RUFIFOOGON</i>	IRGC104624	China	5	22	med/6C
NSFTV472		<i>O. SPONTANEA</i>	IRGC104636	China		22	long/7C
NSFTV477		<i>O. SPONTANEA</i>	IRGC104967	China	9	22	long/7C
NSFTV481		<i>O. NIVARA</i>	IRGC105343	India	9	15	med/6C
NSFTV482		<i>O. RUFIFOOGON</i>	IRGC105349	India	9	15	med/6C
NSFTV483 (483_C2_1_S2)		<i>O. RUFIFOOGON</i>		Thailand	9	15	med/6C
NSFTV487 (487_C2_S2)		<i>O. NIVARA</i>		Sri Lanka	9	15	med/6C
NSFTV490		<i>O. RUFIFOOGON</i>		Japan		24	med/6C
NSFTV492		<i>O. RUFIFOOGON</i>		Japan	9	15	med/6C
NSFTV493		<i>O. NIVARA</i>	IRGC105706	Nepal	9	15	med/6C

Table S2. Germplasm information used for *RAE2* diversity study.

Individuals used for *RAE2* diversity study

Accession ID	Name	Species: subpopulation	Germplasm Repository	Origin	Awn Class (SES)	GC length	RAE2 protein length class
NSFTV494		<i>O. RUFIFOGON</i>	IRGC105711	India		15	med/6C
NSFTV495		<i>O. NIVARA</i>	IRGC105717	Cambodia	9	15	med/6C
NSFTV496		<i>O. RUFIFOGON</i>	IRGC105720	Cambodia		15	med/6C
NSFTV503		<i>O. RUFIFOGON</i>		Thailand	9	15	med/6C
NSFTV505 (505_A1_2_S2)		<i>O. RUFIFOGON</i>		Thailand		21	med/6C
NSFTV508		<i>O. RUFIFOGON</i>	IRGC105890	Bangladesh	0	21	med/6C
NSFTV509		<i>O. RUFIFOGON</i>	IRGC105897	Bangladesh	3	22	long/7C
NSFTV514		<i>O. RUFIFOGON</i>	IRGC105956	Indonesia	9	24	med/6C
NSFTV549		<i>O. RUFIFOGON</i>	IRGC81881	Indonesia	9	24	med/6C
NSFTV553		<i>O. RUFIFOGON</i>	IRGC100926	Japan	9	22	long/7C
NSFTV555 (555_B1_1_S2)					9	15	med/6C
NSFTV592						24	med/6C
NSFTV600			IRGC100187		9	22	long/7C
NSFTV602			IRGC100900		9	15	med/6C
NSFTV605			IRGC100911		9	22	long/7C
NSFTV665		<i>O. RUFIFOGON/O. SATIVA</i>	IRGC100203	Taiwan	0	22	long/7C
NSFTV666		<i>O. RUFIFOGON</i>	IRGC100211	Taiwan	9	15	med/6C
NSFTV669 (669_C2_3_S2)		<i>O. NIVARA</i>		Taiwan	9	21	med/6C
NSFTV673		<i>O. RUFIFOGON</i>	IRGC100647	Taiwan	9	22	med/5C
NSFTV676 (676_A1_1_S2)		<i>O. RUFIFOGON</i>		Taiwan	0	20	short/4C
NSFTV682		<i>O. RUFIFOGON</i>	IRGC100904	Japan	9	24	med/6C
NSFTV701		<i>O. NIVARA/O. RUFIFOGON</i>	IRGC103813	China	9	20	short/4C
NSFTV704		<i>O. RUFIFOGON/O. NIVARA</i>	IRGC103818	China	9	20	short/4C
NSFTV708 (708_A1_2_S2)		<i>O. NIVARA</i>		Bangladesh	9	21	med/6C
NSFTV711		<i>O. NIVARA</i>	IRGC103841	Bangladesh	9	15	med/6C
NSFTV719 (719_A1)		<i>O. NIVARA</i>		France	9	15	med/6C
NSFTV720		<i>O. NIVARA</i>	IRGC104703	France	9	15	med/6C
NSFTV721					9	15	med/6C
NSFTV736 (736_B2_1_S2)					9	24	med/6C
NSFTV743 (743_C1_2_S2T)		<i>O. NIVARA</i>		Nepal	9	15	short/4C
NSFTV751		<i>O. NIVARA</i>	IRGC105895	Bangladesh	9	24	med/6C
NSFTV759 (759_A1_3_S2)		<i>O. RUFIFOGON</i>		Cambodia	9	22	long/7C
NSFTV760		<i>O. NIVARA</i>	IRGC106345	Myanmar	9	21	med/6C
NSFTV762		<i>O. NIVARA</i>		Myanmar	9	15	med/5C
NSFTV765			W1943			24	med/6C
NSFTV767			W1945			24	med/6C
RLS10173		<i>O. barthii</i>	IRGC103912	Tanzania	9	18	med/6C
RLS5584		<i>O. barthii</i>	IRGC104119		9	18	med/6C
RLS10188		<i>O. barthii</i>	IRGC100933		9	18	med/6C
RLS10194		<i>O. barthii</i>	IRGC106303		9	18	med/6C
RLS10128		<i>O. barthii</i>	IRGC101196	Cameroon	9	18	med/6C
RLS10123	WAB 010850	<i>O. barthii</i>	IRGC86524	Chad	9	18	med/6C
RLS10179		<i>O. barthii</i>	IRGC104983	Niger	9	18	med/6C
RLS10183	W0864	<i>O. barthii</i>	IRGC106207	Mali		18	med/6C
RLS10177		<i>O. barthii</i>	IRGC104140	Cameroon	9	18	med/6C
RLS10157		<i>O. barthii</i>	IRGC106291	Mauritania	9	18	med/6C
RLS10190		<i>O. barthii</i>	IRGC100941		9	18	med/6C
RLS10242	DAN MANU (1)	<i>O. glaberrima</i>	TOG5474	Burkina Faso	0	18	med/6C
RLS10239	YAR KARENGESHE	<i>O. glaberrima</i>	TOG5440	Nigeria	0	18	med/6C
RLS10236	TOG5286	<i>O. glaberrima</i>	TOG5286		1	18	med/6C
RLS6183	YANDEV(1)	<i>O. glaberrima</i>	TOG5949	Liberia	1	18	med/6C
RLS10233	ZAKI BIAM-YANDE(WILD)1	<i>O. glaberrima</i>	TOG6193	Nigeria	0	18	med/6C
RLS10245	SHENDAM (WEEDY)1	<i>O. glaberrima</i>	TOG5984	Nigeria	1	18	med/6C
RLS10257	YAR BUTUKA	<i>O. glaberrima</i>	TOG5467	Nigeria	1	18	med/6C
RLS10253	DAN MAIWUYA (6)	<i>O. glaberrima</i>	TOG5390	Nigeria	0	18	med/6C
RLS10262	NEW AYOMA LOCAL (2)	<i>O. glaberrima</i>	TOG7402	Ghana	5	18	med/6C
RLS10247	SHAWHON (2)	<i>O. glaberrima</i>	TOG5747	Liberia	1	18	med/6C
RLS10264	ACC 100982	<i>O. glaberrima</i>	TOG6211	Guinea Bissau	0	18	med/6C
RLS10265	QUE (2)	<i>O. glaberrima</i>	TOG5815	Liberia	0	18	med/6C

Table S3. Seven different length polymorphisms in the GC-repeat region of *RAE2*.

	nt length	***protein length class	Cys no.	wild Asian	cult. Asian	wild African	cult. African
African	18	medium				10	11
	15	medium	5C/ 6C	25*			
	21	medium		9			
Asian	24	medium		9	4		
	20	short	4C	5	17		
	23	short		1			
	22	long	7C	12**	20		

*one accession showed short RAE2 protein length (4C) because of the 1 bp insertion in apart from GC-rich region

**one accession showed medium RAE2 protein length (5C) because of the 1 bp insertion in apart from GC-rich region

***the classification of protein length class is described in Fig. S7B legend

Table S4. Germplasm information used for *RAE2* selective sweep analysis.

Accession ID	Name	Species: subpopulation	Germplasm Repository	Origin	RAE2 protein length class
NSFTV1	Agostano	<i>O. sativa: temperate japonica</i>	IRGC126380		4C/short
NSFTV104	Mansaku	<i>O. sativa: temperate japonica</i>	IRGC117811		6C/med
NSFTV108	Moroberekan	<i>O. sativa: tropical japonica</i>	IRGC117621		4C/short
NSFTV110	Mudgo	<i>O. sativa: indica</i>	IRGC117818		7C/long
NSFTV154	Ta_Hung_Ku	<i>O. sativa: temperate japonica</i>	IRGC117904		6C/med
NSFTV16	Bico_Branco	<i>O. sativa: aromatic</i>	IRGC117658		4C/short
NSFTV163	Taducan	<i>O. sativa: indica</i>	IRGC117906		7C/long
NSFTV165	Trembese	<i>O. sativa: tropical japonica</i>	IRGC117921		7C/long
NSFTV173	Nipponbare	<i>O. sativa: temperate japonica</i>			4C/short
NSFTV174	Azucena	<i>O. sativa: tropical japonica</i>			7C/long
NSFTV18	BJ1	<i>O. sativa: aus</i>	IRGC117661		4C/short
NSFTV19	Black_Gora	<i>O. sativa: aus</i>	IRGC117662		4C/short
NSFTV207	Sigadis	<i>O. sativa: indica</i>	IRGC117889		7C/long
NSFTV226	IRAT_44	<i>O. sativa: tropical japonica</i>	IRGC117762		7C/long
NSFTV23	Canella_De_Ferro	<i>O. sativa: tropical japonica</i>	IRGC117675		7C/long
NSFTV248	Caucasica	<i>O. sativa: temperate japonica</i>	IRGC117677		4C/short
NSFTV268	Vavilovi	<i>O. sativa: temperate japonica</i>	IRGC117928		4C/short
NSFTV28	Champa_Tong_54	<i>O. sativa: aus</i>	IRGC117680		4C/short
NSFTV29	Chau	<i>O. sativa: indica</i>	IRGC117682		7C/long
NSFTV317	DJ123	<i>O. sativa: aus</i>	IRGC117711		4C/short
NSFTV336	Paung_Malaung	<i>O. sativa: aus</i>	IRGC117847		4C/short
NSFTV338	Sitpwa	<i>O. sativa: temperate japonica</i>	IRGC117892		4C/short
NSFTV341	Shirkati	<i>O. sativa: aus</i>	IRGC117885		4C/short
NSFTV369	Sathi	<i>O. sativa: aus</i>	IRGC117878		4C/short
NSFTV378	Kalubala_Vee	<i>O. sativa: aus</i>	IRGC117774		4C/short
NSFTV397	Cybonnet	<i>O. sativa: tropical japonica</i>	IRGC117699		4C/short
NSFTV402		<i>O. spontanea</i>	IRGC80539	India	4C/short
NSFTV410		<i>O. nivara</i>	IRGC80759	Myanmar	6C/med
NSFTV413		<i>O. nivara</i>	IRGC81850	India	6C/med
NSFTV415		<i>O. spontanea</i>	IRGC81909	India	6C/med
NSFTV416		<i>O. spontanea</i>	IRGC81970	Thailand	7C/long
NSFTV422		<i>O. rufipogon</i>		Vietnam	6C/med
NSFTV427		<i>O. rufipogon</i>		China	6C/med
NSFTV43	Dee_Geo_Woo_Gen	<i>O. sativa: indica</i>	IRGC117705		7C/long
NSFTV431		<i>O. rufipogon</i>	IRGC82992	China	6C/med
NSFTV432		<i>O. rufipogon</i>		Thailand	7C/long
NSFTV433		<i>O. rufipogon</i>	IRGC83795	India	4C/short
NSFTV435		<i>O. rufipogon</i>	IRGC86448	Thailand	6C/med
NSFTV438 (438_B2_1_S2)		<i>O. rufipogon</i>		India	7C/long
NSFTV443		<i>O. nivara</i>	IRGC93183	Nepal	6C/med
NSFTV444		<i>O. nivara</i>	IRGC93188	Nepal	6C/med
NSFTV446		<i>O. spontanea</i>	IRGC93224	Nepal	6C/med
NSFTV450		<i>O. nivara</i>	IRGC100916	China	6C/med
NSFTV453		<i>O. rufipogon</i>	IRGC103404	Bangladesh	7C/long
NSFTV457 (457_B3_1_S2)		<i>O. nivara</i>		Bangladesh	6C/med
NSFTV461 (461_A1_1_S2)		<i>O. rufipogon</i>		China	4C/short
NSFTV467		<i>O. RUFIPOGON</i>	IRGC104624	China	6C/med
NSFTV472		<i>O. SPONTANEA</i>	IRGC104636	China	7C/long
NSFTV477		<i>O. SPONTANEA</i>	IRGC104967	China	7C/long
NSFTV481		<i>O. NIVARA</i>	IRGC105343	India	6C/med
NSFTV482		<i>O. RUFIPOGON</i>	IRGC105349	India	6C/med
NSFTV483 (483_C2_1_S2)		<i>O. RUFIPOGON</i>		Thailand	6C/med
NSFTV487 (487_C2_S2)		<i>O. NIVARA</i>		Sri Lanka	6C/med
NSFTV490		<i>O. RUFIPOGON</i>		Japan	6C/med
NSFTV492		<i>O. RUFIPOGON</i>		Japan	6C/med
NSFTV493		<i>O. NIVARA</i>	IRGC105706	Nepal	6C/med
NSFTV494		<i>O. RUFIPOGON</i>	IRGC105711	India	6C/med
NSFTV495		<i>O. NIVARA</i>	IRGC105717	Cambodia	6C/med
NSFTV496		<i>O. RUFIPOGON</i>	IRGC105720	Cambodia	6C/med
NSFTV503		<i>O. RUFIPOGON</i>		Thailand	6C/med

Table S4. Germplasm information used for *RAE2* selective sweep analysis.

Accession ID	Name	Species: subpopulation	Germplasm Repository	Origin	RAE2 protein length class
NSFTV505 (505_A1_2_S2)		<i>O. RUFIOGON</i>		Thailand	6C/med
NSFTV508		<i>O. RUFIOGON</i>	IRGC105890	Bangladesh	6C/med
NSFTV509		<i>O. RUFIOGON</i>	IRGC105897	Bangladesh	7C/long
NSFTV51	Early_Wataribune	<i>O. sativa: temperate japonica</i>	IRGC117727		4C/short
NSFTV514		<i>O. RUFIOGON</i>	IRGC105956	Indonesia	6C/med
NSFTV549		<i>O. RUFIOGON</i>	IRGC81881	Indonesia	6C/med
NSFTV553		<i>O. RUFIOGON</i>	IRGC100926	Japan	7C/long
NSFTV555 (555_B1_1_S2)					6C/med
NSFTV56	Geumobyeyo	<i>O. sativa: temperate japonica</i>	IRGC117612		4C/short
NSFTV57	Gharib	<i>O. sativa: indica</i>	IRGC117739		7C/long
NSFTV592					6C/med
NSFTV600			IRGC100187		7C/long
NSFTV602			IRGC100900		6C/med
NSFTV605			IRGC100911		7C/long
NSFTV612	IR64	<i>O. sativa: indica</i>			7C/long
NSFTV620	Jasmine85	<i>O. sativa: indica</i>	IRGC125597		7C/long
NSFTV628	Jefferson	<i>O. sativa: tropical japonica</i>	IRGC126385		4C/short
NSFTV630	Saber	<i>O. sativa: tropical japonica</i>	IRGC126393		4C/short
RLS672	Minghui_63	<i>O. sativa: indica</i>	IRGC117271		7C/long
NSFTV665		<i>O. RUFIOGON/O. SATIVA</i>	IRGC100203	Taiwan	7C/long
NSFTV666		<i>O. RUFIOGON</i>	IRGC100211	Taiwan	6C/med
NSFTV669 (669_C2_3_S2)		<i>O. NIVARA</i>		Taiwan	6C/med
NSFTV673		<i>O. RUFIOGON</i>	IRGC100647	Taiwan	5C/med
NSFTV676 (676_A1_1_S2)		<i>O. RUFIOGON</i>		Taiwan	4C/short
NSFTV682		<i>O. RUFIOGON</i>	IRGC100904	Japan	6C/med
NSFTV7	Arias	<i>O. sativa: tropical japonica</i>	IRGC126381		7C/long
NSFTV701		<i>O. NIVARA/O. RUFIOGON</i>	IRGC103813	China	4C/short
NSFTV704		<i>O. RUFIOGON/O. NIVARA</i>	IRGC103818	China	4C/short
NSFTV708 (708_A1_2_S2)		<i>O. NIVARA</i>		Bangladesh	6C/med
NSFTV711		<i>O. NIVARA</i>	IRGC103841	Bangladesh	6C/med
NSFTV719 (719_A1)		<i>O. NIVARA</i>		France	6C/med
NSFTV720		<i>O. NIVARA</i>	IRGC104703	France	6C/med
NSFTV721					6C/med
NSFTV736 (736_B2_1_S2)					6C/med
NSFTV743 (743_C1_2_S2T)		<i>O. NIVARA</i>		Nepal	4C/short
NSFTV751		<i>O. NIVARA</i>	IRGC105895	Bangladesh	6C/med
NSFTV759 (759_A1_3_S2)		<i>O. RUFIOGON</i>		Cambodia	7C/long
NSFTV760		<i>O. NIVARA</i>	IRGC106345	Myanmar	6C/med
NSFTV762		<i>O. NIVARA</i>		Myanmar	5C/med
NSFTV765			W1943		6C/med
NSFTV767			W1945		6C/med
NSFTV81	Kalamkati	<i>O. sativa: aus</i>	IRGC117773		4C/short
NSFTV84	Kaniranga	<i>O. sativa: tropical japonica</i>	IRGC117776		7C/long
NSFTV85	Kasalath	<i>O. sativa: aus</i>	IRGC117617		4C/short
RLS29440	KUI_SALI	<i>O. sativa: aromatic</i>			4C/short
RLS49428	JC111	<i>O. sativa: aromatic</i>			4C/short
RLS29427	JC101	<i>O. sativa: aromatic</i>			4C/short
RLS5667	ARC_13523	<i>O. sativa: aromatic</i>			4C/short
RLS5669	Sathi_Basmati	<i>O. sativa: aromatic</i>			4C/short
RLS5665	Ambemohar	<i>O. sativa: aromatic</i>			4C/short
RLS29412	Basmati	<i>O. sativa: aromatic</i>			4C/short
RLS5670	Taraori_Basmati	<i>O. sativa: aromatic</i>			4C/short
RLS6303	Basmati_370	<i>O. sativa: aromatic</i>			4C/short
RLS29421	Sadri_Belyi	<i>O. sativa: aromatic</i>			4C/short
RLS29461	X9524	<i>O. sativa: aus</i>			4C/short
RLS29443	Khao_Gaew	<i>O. sativa: aus</i>			4C/short
RLS29426	Gie_57	<i>O. sativa: aus</i>			4C/short
RLS29437	BADAL89	<i>O. sativa: aus</i>			4C/short
RLS29464	Jhona349	<i>O. sativa: aus</i>			4C/short
RLS367	Chati_Kamma_Nangarhar	<i>O. sativa: aus</i>			4C/short

Table S4. Germplasm information used for *RAE2* selective sweep analysis.

Accession ID	Name	Species: subpopulation	Germplasm Repository	Origin	RAE2 protein length class
RLS29433	TD2	<i>O. sativa: indica</i>			7C/long
RLS29419	Leung_Prataew	<i>O. sativa: indica</i>			7C/long
RLS29431	Popot_165	<i>O. sativa: indica</i>			7C/long
RLS29418	Guan.Yin.Tsan	<i>O. sativa: indica</i>			7C/long
RLS29429	JC91	<i>O. sativa: indica</i>			7C/long
RLS5364	CO39	<i>O. sativa: indica</i>			7C/long
RLS930	Short_Grain	<i>O. sativa: indica</i>			7C/long
RLS29463	X9311	<i>O. sativa: indica</i>			7C/long
RLS460	IR8	<i>O. sativa: indica</i>			7C/long
RLS5316	Taichungsien17	<i>O. sativa: indica</i>			7C/long
RLS5317	Tainungsien20	<i>O. sativa: indica</i>			7C/long

Table S5. *RAE2* variants across the five subpopulations of *O. sativa*.

Protein type	<i>tej</i>	<i>trj</i>	<i>aro</i>	<i>aus</i>	<i>ind</i>
Short (dysfunc.)	0.33	0.33	1.00	1.00	
Medium (func.)	0.33	0.18			
Long (dysfunc.)	0.33	0.55			1.00
n	6	11	6	6	12

Table S6. The list of SP-8 type protease expressed in the spikelet.

No.	Locus IDs	Chromosome location	Accession	FeatureNum (Link to graph)
1	LOC_Os01g17160	chr01	AK119444	36860
2	LOC_Os01g50680	chr01	-	5682
3	LOC_Os01g52750	chr01	AK108195	17131
4	LOC_Os01g56320	chr01	AK070376	37342
5	LOC_Os01g58240	chr01	AK109067	14086
6	LOC_Os01g58260	chr01	AF200467	30347
7	LOC_Os01g58270	chr01	-	1768
8	LOC_Os01g58280	chr01	AK066488	6645
9	LOC_Os01g58290	chr01	AK100351	18445
10	LOC_Os01g6485	chr01	-	-
11	LOC_Os01g64860	chr01	AK062271	19280
12	LOC_Os02g10520	chr02	AK120287	30014
			AK106394	40115
13	LOC_Os02g16940	chr02	-	8596
14	LOC_Os02g17000	chr02	AK110825	25369
15	LOC_Os02g17060	chr02	-	-
16	LOC_Os02g17080	chr02	-	-
17	LOC_Os02g17090	chr02	AK072092	42776
18	LOC_Os02g17150	chr02	-	35042
19	LOC_Os02g44520	chr02	AK103515	33598
			AK067099	34821
			AK066478	41274
20	LOC_Os02g44590	chr02	AB037371	19329
			AK072929	20201
			AK100551	21433
21	LOC_Os02g53850	chr02	-	16087
22	LOC_Os02g53860	chr02	AK106527	5667
			AK121728	42860
23	LOC_Os02g53910	chr02	-	-
24	LOC_Os02g53970	chr02	AK070669	30949
25	LOC_Os03g02750	chr03	AK069220	12951
26	LOC_Os03g04950	chr03	-	-
27	LOC_Os03g06290	chr03	AK071242	43120
28	LOC_Os03g13930	chr03	-	39736
29	LOC_Os03g31630	chr03	-	37303
30	LOC_Os03g40830	chr03	AK105749	29457
31	LOC_Os03g55350	chr03	AK101646	22817
			AK103255	31812
32	LOC_Os04g02960	chr04	CI260116	1713
33	LOC_Os04g02980	chr04	-	4178
34	LOC_Os04g03060	chr04	-	16547
35	LOC_Os04g03100	chr04	-	26338
36	LOC_Os04g03710	chr04	-	-
37	LOC_Os04g03800	chr04	-	-
38	LOC_Os04g03810	chr04	AK062269	23506
39	LOC_Os04g03850	chr04	-	-
40	LOC_Os04g10360	chr04	CB653384	32192
41	LOC_Os04g35140	chr04	AK105112	5780
42	LOC_Os04g45960	chr04	AK106823	12015
			AY644644	15417
			AY683198	21133
43	LOC_Os04g47150	chr04	AK100861	1851
44	LOC_Os04g47160	chr04	-	23449
45	LOC_Os04g48420	chr04	-	-
46	LOC_Os05g30580	chr05	AK064686	5180
47	LOC_Os05g36010	chr05	AK067138	31360
48	LOC_Os06g06800	chr06	-	-
49	LOC_Os06g06810	chr06	AK071415	24262
50	LOC_Os06g40700	chr06	AK109185	4261
51	LOC_Os06g41880	chr06	-	39167
52	LOC_Os06g48650	chr06	AK102835	3665
53	LOC_Os07g39020	chr07	AK107610	24754
54	LOC_Os07g48650	chr07	AK119348	41308
55	LOC_Os08g23740	chr08	-	-
56	LOC_Os08g35090	chr08	CI043104	38603
57	LOC_Os09g26920	chr09	CI383807	8743
58	LOC_Os09g30250	chr09	CI269495	4777
59	LOC_Os09g36110	chr09	-	16075
60	LOC_Os10g25450	chr10	CI191448	30831
61	LOC_Os10g38080	chr10	AK069238	21727
62	LOC_Os11g15520	chr11	AK110921	33825
63	LOC_Os12g23980	chr12	-	23981

Table S7. Primers used in this study.

Purpose	name	Sequence (5' → 3')
Linkage mapping of RAE2	8KG23941	CACGCTTGTAAAGGCTGAGTT ATTCCGTATCCGAAAACCTC
	8KG23994	TGGAACAACGTGAGATTGTC GTTCTGATCAGATTGTTGC
	8KG23999	CATCCATCAACATGTCGTCCG CGCCATGTATAGTGTGATTCCG
	8KG24021	TATCCTTCTTGGGTTCTTGC TGAATGTGGTGCATTTTCATC
BAC screening	pk31	GCACCTCAGCCTGGTTTCAAG
	pk32	GTAGTAGTTTGGTTGTTCTCTTGC
RAE2 promoter sequencing	KU42	CCAAGATGACAGCATGCTACTG
	KU43	CCAATTCTTTGTAACAAAGGGTAG
RAE2 coding region cloning	KU32	CACCATGAGGACGGCGGCCACGCCGCT
	KU35	TCAGGGGTCGAACAGGCG
RAE2 RNAi construct	RNAi-F	CACCGATAGATTCCGTGTAATAT
	RNAi-R	ATATTACCGGAATCTATC
qRT-PCR of RAE2	KU37	ATTTTGACCAGACCACCTCG
	KU38	CGCCCAGCTACTTATACCCA
qRT-PCR of ACT1	ACT1 RT-f	GGATCCATCTTGGCATCTCTCA
	ACT1 RT-r	GGGCCAGACTCGTCTACTC
qRT-PCR of UBQ5	UBQ5 RT-f	AAACCCTAACGGGGAAGACCATAA
	UBQ5 RT-r	CCACAGTAATGGCGATCAAATGA
in situ probe of RAE2	in situ-f	AGCTTCTTGGTAGGCGAGGTGT
	in situ-r	GAAGAAGACGGCGAGGAGGA
3xFLAG primer	KU71	CGACCATGGGATTACAAGGATGACGACGATAAGGACTATAAGGACGATGATGACAAGGATTACAAAGATG
	KU72	CGACCATGGTTTATCGTCATCATCTTTGTAATCCTTGTTCATCATCGTCCTTATAGTCCTTATCGTCGTCA
pACT::RAE2-3xFLAG	KU73	ATGTCTAGAATGAGGACGGCGGCCACGCC
	KU75	CTAAGCTTACGGGGTCGAACAGGCGGTTCGC
recombinant RAE2 pro-peptide	PRO-f	CCGATATCAGGCTCCCTCCTCCTCGC
	PRO-r	GGTGCTCGAGCTACCATGGTTTATCGTCAT
recombinant RAE2 mature peptide	MA-f	GATCGGATCCCGGCTGGGGTCGAGCCCCGC
	MA-r	CCCAAGCTTTCAGGGGTCGAAC
Alanine substituted construct #1	KU132	GAGGAGAAGGTGCGGGCGGGCGGGCGGGCGGAGCTGCTACAGCAAGTGC
	KU133	CCCGTAGCACTTGTGTAGCAGCTCGCCGCCGCCGCCGCCGCCGCCGCCACCTTCTC
Alanine substituted construct #2	KU111	GGGGAGGAGGAGGCGGCGCGCTGGGGTTCG
	KU112	CAGCCGCGCCGCTCCTCCTCCCCAGCCAC
Alanine substituted construct #3	KU113	GAGGAGGAGAAGGCGGCGCTGGGGTTCGAGC
	KU114	CCCCAGCGCCGCTTCTCCTCCTCCCCAGC
Alanine substituted construct #4	KU121	GCTGTAGCAGCTCGCCGCGCTCGACCCAG
	KU122	CTCGACCCAGCCGACCCGCCGCCGCCCTCCCCAGC
SLP1 cloning	KU127	CACCATGCAGACTTATGTGATCGTCTTTG
	KU139	GGAATTCCCTACCCGAGGTGCTCTTG
semi qRT-PCR of SLP1	KU144	CGTGTCCCCTACAACATAATGTCC
	KU146	GATCTTGCCGCTGTGCTTGTGTC
in vitro translation of SLP1	FLAG-f	CCAGCAGGGAGGTACTATGCAGACTTATGTGATCGT
	FLAG-r	CCTTATGGCCGATCCAAGAGCTCTTTTTTTTTTTTACCCGAGGTGCTCTTG